Learning and Mining from Personal Digital Archives

Na Li, M.Sc., B.Sc.

A dissertation submitted in fulfilment of the requirements for the award of

Doctor of Philosophy (Ph.D.)

to the

Dublin City University

School of Computing

Supervisors:

Assoc. Prof. Martin Crane

Prof. Heather J. Ruskin

Assoc. Prof. Cathal Gurrin

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:

ID No: 12211574

Date:

Acknowledgements

Studying part-time towards a Ph.D degree was challenging, I am remarkable lucky to have met so many wonderful people who accompanied me on this journey.

Firstly, I would like to thank my supervisors for their great guidance and support which made this thesis possible. I have learned a lot from them from various perspectives. I especially thank my supervisor Prof. Heather J. Ruskin for her patience, in-depth comments and invaluable guidance throughout my studies. I thank my supervisor Dr. Martin Crane for always being a good listener, for giving sincere advice and for providing help for everything. I also thank my supervisor Dr. Cathal Gurrin for all his inputs, support and opportunities. It was my great honour to be able to work with them. Thank you also my two examiners, Prof. Petia Radeva (U. Barcelona), Dr. John McKenna and my viva chair Prof. Andy Way for an enjoyable viva experience.

I would particularly like to express my gratitude to my former employers Prof. Turlough Downes (Sci-Sym), Mr. Ray Walshe (CloudCORE), Prof. Theo Lynn (IC4) and Prof. Marcus Greferath (UCD) for their support in the past years. Thanks to them, I was able to self-fund my PhD and did not have to face financial worries during my studies. I truly appreciate this.

I also would like to thank School of Computing for various generous financial support and many current and former colleagues who provided advice, support and help over the course of this research. I would particularly like to thank Dr. Irina Balaur, Dr. Marija Bezbradica, Dr. Alistair Sutherland, Dr. Liam Tuohey and Mrs. Patricia Lacey for all their help and support.

I enjoyed my PhD studies thanks to the friendship and support from my Chinese friends whom I have met in DCU, Jiang, Wei, Zhenxing, Lulu, Dian, Yun, Yuki, Lei, Lily and Huanhuan. Especially Feiyan who started his PhD with me together and who shared all his knowledge and had helpful discussions with me. I thank you all.

I would also like to thank my beloved parents and my uncle for all their constant love, encouragement and support, to make my dreams do come true. Thank you to my cousins, my nephew, my niece from China and my family members from Germany for their constant emotional support.

Finally, I would like to thank my wonderful husband and soul-mate, Frank, for all his love. He always preserved his confidence in my ability to complete my PhD and provided huge support for my career. I am very fortunate to have him by my side to share all happiness and sadness during my entire PhD studies. He always understanding. Ich liebe dich über alles!

Contents

Acknowledgements **i**

List of Figures **v**

List of Tables **ix**

Abstract **x**

1 Introduction **1**

1.1 Motivation 1

1.2 Objectives 5

1.3 Outline of Thesis 6

2 Background **8**

2.1 Introduction 8

2.2 History and Applications of Lifelogging 9

2.3 State-of-the-Art in Visual Lifelogging 12

2.4 Complex Systems 20

2.5 Time Series Analysis 22

2.5.1 Stationary and Non-Stationary Time Series 23

2.5.2 Univariate and Multivariate Time Series 23

2.5.3 Detrended Fluctuation Analysis 24

2.5.4 Cross-Correlation Analysis and Noise Reduction 25
2.5.5 Wavelet Analysis 27
2.5.6 Time Series Motifs 29
2.6 Data Sets Used in this Thesis 31
2.7 Summary 40

3 Multiscaled Cross-Correlation Dynamics 42
3.1 Introduction 42
3.2 Methods 42
3.2.1 Detrended Fluctuation Analysis 43
3.2.2 Correlation Dynamics 45
3.2.3 Wavelet Multiscale Analysis 47
3.3 Results 52
3.3.1 Detrended Fluctuation Analysis 53
3.3.2 Dynamics of the largest Eigenvalue for different sliding window sizes 54
3.3.3 Wavelet Multiscale Anlaysis 55
3.3.4 Evaluation 64
3.4 Summary 67

4 Random Matrix Theory 69
4.1 Introduction 69
4.2 Methods 70
4.2.1 Random Matrix Theory 70
4.2.2 Eigenvector Analysis 72
4.3 Results 73

4.3.1 Statistics of Correlation Coefficients 73
4.3.2 Eigenvalue Analysis 74
4.3.3 Eigenvector Analysis 77
4.4 Summary 84

5 Finding Motifs in Large Personal Lifelogs **86**
5.1 Introduction 86
5.2 Methods 87
5.2.1 Symbolic Aggregate approXimation (SAX) 87
5.2.2 Estimating Extracted Motif Candidate Based on MDL Principle 90
5.3 Results 95
5.4 Summary 106

6 Conclusions and Future Work **108**
6.1 Summary 108
6.2 Future Work 111

Bibliography **116**

List of Figures

2.1 The **Memex**. Source: *Bush & Think (1945)* 10

2.2 Overview of MyLifeBits. Source: *Gemmell et al. (2006)* 11

2.3 Examples of images taken by author-worn SenseCam. 13

2.4 A range of the popular wearable cameras. Source: *Streams (2012), Michael et al. (2018)* 15

2.5 Segmenting a day of SenseCam images into distinct events. Source: *Doherty (2009)* 19

2.6 The wavelet transform partitioning of frequency and time plane. Source: *Gallegati (2008)* 28

2.7 An astronomical time series contains 3 near identical subsequences. Source: *Lin et al. (2002)* 30

2.8 The SAX motif discovery technique. The different letters represent repeated motifs or patterns. Source: *Lin et al. (2007)* . . 30

2.9 Autographer Wearable Camera and Sample; an image from the NTCIR12-Lifelog Test Collection 38

2.10 Outline of the Research Methods covered by this thesis 41

3.1 Flowchart for a three-level DWT. Source from: *Seo et al. (2017)* 48

3.2 Flowchart for a three-level MODWT. Source from: *Seo et al. (2017)* 49

3.3 Plot of log $F(n)$ vs log n , where n = box size, ranging from 10 to 1000 for the AUTHOR6DAYS data set (D_1). 53
3.4 The largest Eigenvalue Distribution using a sliding window of 50 Images(a), 100 Images(b), 200 Images(c) and 400 Images(d) of the AUTHOR6DAYS data set (D_1). 54
3.5 Examples of wavelet filters (a) Haar (b) Daublets (c) Coiflets (d) Symmlets. Source:*(Zubal' 2015)* 56
3.6 Heatmap diagram of the largest Eigenvalue λ_1 dynamics across 9 wavelet scales of the AUTHOR6DAYS data set (D_1). Scales 1 *(a)* to 9 *(i)* correspond to time periods of 1-2, 2-4, 4-8, 8-16, 16-32, 32-64, 64-128, 128-256 and 256-512 minutes, respectively 57
3.7 The Largest Eigenvalue λ_1 *(red)* and ratio λ_1/λ_2 *(black)* dynamics across 9 wavelet scales of the AUTHOR6DAYS data set (D_1). *(a-c)* are for scales 1 to 3 (1-2, 2-4 and 4-8 minutes periods, respectively), *(d-f)* are for scales 4 to 6 (8-16, 16-32 and 32-64 minute periods, respectively) and *(g-i)* are for scales 7 to 9 (64-128, 128-256 and 256-512 minute periods, respectively.) . . 59

4.1 Correlation Coefficient Distribution of the AUTHORTYPICAL data set (D_2) for the Cross-Correlation Matrix $\mathbf{C}$ for SenseCam data (black) and Random Matrix $\mathbf{R}$ (red). 74
4.2 Eigenvalue Distribution of the AUTHORTYPICAL data set (D_2) for the Correlation Matrix $\mathbf{C}$ for SenseCam data: (a) Full and (b) Partial spectral distribution. 76

4.3 Comparison of Eigenvector Components of the AUTHORTYPICAL data set (D_2), (a) largest Eigenvector, (b) second largest Eigenvector, (c) third largest Eigenvector and (d) Eigenvector from the bulk 78

4.4 Probability distribution P of the Cross-Correlation coefficients of the AUTHORTYPICAL data set (D_2) for data before (black) and after (red) removing the effect of the largest eigenvalue by linear regression method 79

4.5 Inverse Participation Ratio (IPR) of the AUTHORTYPICAL data set (D_2) as function of eigenvalue λ for the (a) Random Matrix R and (b) Cross-Correlation Matrix C 81

4.6 Distribution of eigenvector components: (a) largest remaining eigenvalue (b) 2nd largest and (c) 3rd largest remaining eigenvalue for the AUTHORTYPICAL data set (D_2) 83

5.1 Example of time series transformation into SAX symbols. Here, n=112, w=16, a=4. The time series is mapped to the PAA symbols **bcbbaaaaccbdbbcb.** 89

5.2 Example of 'behaviour symbol' assignment for pattern order of PAA symbols. In this, **A=bacc, B=ccbc and C=bcdc.** . . . 89

5.3 Calculation of the MDL pattern algorithm 94

5.4 The largest eigenvalue λ_1 (*blue*) dynamics with the curve generated for the PAA (*red*). (a) Wavelet scale 1, (b) scale 2, (c) scale 3 and (d) scale 4 for the AIHSWEDS data set (D_3). 96

5.5 Example of Motifs extracted based on MDL principle from different wavelet scales. (a) Wavelet scale 1, (b) Wavelet scale 2, (c) Wavelet scale 3 and (d) Wavelet scale 4 for the AIHSWEDS data set (D_3). 98

5.6 Heatmap diagram showing the dynamics of the largest Eigenvalue λ_1 across 9 wavelet scales. Scales 1 to 9 correspond respectively to periods of 1-2 , 2-4, 4-8, 8-16, 16-32, 32-64, 64-128, 128-256 and 256-512 minutes, respectively, for the NTCIR-12 (3 USERS) data set (D_4). 102

List of Tables

2.1 The AUTHOR6DAYS data set (D_1) 33

2.2 The AUTHORTYPICAL data set (D_2) 34

2.3 AIHS Data Set Overall Summary 36

2.4 The AIHSWEDS data set (D_3) 37

2.5 The NTCIR-12 (3 USERS) data set (D_4) 39

3.1 Precision, Recall and F_1 measures for MODWT method 66

5.1 Summary of the Notation of the SAX Algorithm 91

5.2 Summary of Ground Truth of the NTCIR-12 (3 USERS) data set (D_4) 100

5.3 Motif: 'Working in front of Computer' Event 104

5.4 The accuracy of different motifs 105

Abstract

Given the explosion of new sensing technologies, data storage has become significantly cheaper and consequently, people increasingly rely on wearable devices to create personal digital archives. *Lifelogging* is the act of recording aspects of life in digital format for a variety of purposes such as aiding human memory, analysing human lifestyle and diet monitoring. In this dissertation we are concerned with *Visual Lifelogging*, a form of lifelogging based on the passive capture of photographs by a wearable camera. Cameras, such as Microsoft's SenseCam can record up to 4,000 images per day as well as logging data from several incorporated sensors. Considering the volume, complexity and heterogeneous nature of such data collections, it is a significant challenge to interpret and extract knowledge for the practical use of lifeloggers and others.

In this dissertation, time series analysis methods have been used to identify and extract useful information from temporal lifelogging images data, without benefit of prior knowledge. We focus, in particular, on three fundamental topics: noise reduction, structure and characterization of the raw data; the detection of multi-scale patterns; and the mining of important, previously unknown repeated patterns in the time series of lifelog image data.

Firstly, we show that Detrended Fluctuation Analysis (DFA) highlights the feature of very high correlation in lifelogging image collections. Secondly, we show that study of equal-time Cross-Correlation Matrix demonstrates atypical or non-stationary characteristics in these images. Next, noise reduction in the Cross-Correlation Matrix is addressed by Random Matrix Theory (RMT) before Wavelet multiscaling is used to characterize the 'most important' or 'unusual' events through analysis of the associated dynamics of the eigenspectrum. A motif discovery technique is explored for detection of recurring and recognizable episodes of an individual's image data. Finally, we apply these motif discovery techniques to two known lifelog data collections, All I Have Seen (AIHS) and NTCIR-12 Lifelog, in order to examine multivariate recurrent patterns of multiple-lifelogging users.

Chapter 1

Introduction

1.1 Motivation

Many people have recently developed the habit of keeping certain documents or materials (e.g. photographs, postcards, letters, bills, etc) throughout their day-to-day lives. Such ways of storing collections can be considered as personal archives. With the ubiquity of personal computing devices and advances in storage technology, the process or action of storing increasing amounts of personal information in digital format, is referred to as *Personal Digital Archives* or *Personal Informatics* (Waters & Garrett 1996, Lindley et al. 2013, Rooksby et al. 2014). Today, individuals can choose from a variety of tools and services to achieve this. Examples include messaging systems (Email, WhatsApp, SMS...), social network activities (Facebook, Twitter, Linkedin...), data shared and stored in the information cloud (YouTube, Pinterest, Instagram..) or posted directly on Web blogs (WordPress, Blogger, Tumblr...). In particular, as these activities are directly connected to the rise of the *Internet of Things* (IoTs), this topic has received a lot of attention over the last years

(Atzori et al. 2010, Gubbi et al. 2013, Sarkar et al. 2018).

With the proliferation of new sensing technologies has come significantly cheaper hard-drive or cloud storage and consequently, this has led individuals increasingly to rely on such wearable devices to create their personal digital archives. There is also a growing number of people interested in constant measurement of the performance level of their bodies, where these are using a variety of types of equipment to collect and store the data. These individuals can be considered as part of the *Quantified Self* (QS) movement[1] that uses instruments to record numerical data on all aspects of our lives: input (food consumed, surrounding air-quality), states (mood, arousal, blood oxygen levels), and performance (mental, physical). Such self-monitoring and self-sensing, which combines wearable sensors and wearable computing, is also sometimes referred to as *Lifelogging*. Lifelogging information shares features (volume, variety, velocity and veracity) usually identified with so called *Big Data* (Gurrin et al. 2014).

Lifelogging has been around for a long time, but has only become popular in recent years. Consequently a generally accepted definition of lifelogging has yet to crystallize. Lifelogging can however be defined as "a form of pervasive computing consisting of a unified digital record of the totality of an individual's experiences, captured multimodally through digital sensors and stored permanently as a personal multimedia archive" (Dodge & Kitchin 2007). Recently, Gurrin et al. (2014) have argued that "lifelogging represents a phenomenon whereby people can digitally record their own daily lives in varying amounts of detail, for a variety of purposes". "Lifelogging" is considered as a process of

[1]http://quantifiedself.com

automatically, passively and digitally recording aspects of our life experience, creating a comprehensive archive of an individual's quotidian existence. Lifelogs can consist of such unstructured data as heterogeneous sensor data such as images, GPS coordinates, WiFi streams, accelerometer and light level measurements amongst others. Lifelogging can take many forms, such as capturing everything seen through wearable cameras, detection of people met (through bluetooth devices), identifying places visited (through GPS), calculating distances or speed (by accelerometer readings), etc. One vision of lifelogging is that the user never forgets anything since everything is being tracked, photographed and recorded. Lifelogging can provide useful information for the lifeloggers in different domains such as lifestyle analysis (Doherty et al. 2011, Bolaños et al. 2013), diet monitoring (Reddy et al. 2007), analysis of Activities of Daily Living (ADLs) (Cartas et al. 2017), recognition of physical activities (Zhang et al. 2010), memory reinforcement (Spector et al. 2003, Sellen et al. 2007, Harvey et al. 2016) and so on.

A special case of lifelogging is *Visual Lifelogging* (also called *Egocentric Vision* (Doherty et al. 2013, Betancourt et al. 2015, Bolaños et al. 2017)), where lifeloggers wear a camera mounted on the head (Hori & Aizawa 2003, Mann et al. 2005) or chest (Blum et al. 2006, Sellen et al. 2007), that captures personal activities through the medium of images or video. Despite its relative novelty, visual lifelogging is gaining popularity due to projects such as the Microsoft SenseCam camera (Hodges et al. 2006). The SenseCam is a small, lightweight wearable device that automatically captures a wearer's every moment as a series of images and sensor readings. Normally, the device captures an image at the rate of one every 30 seconds and collects about 4,000

images in a typical day. Research shows that the SenseCam can be an effective memory-aid device (Spector et al. 2003, Sellen et al. 2007, Harper et al. 2007, 2008, Harvey et al. 2016), as it helps to improve retention of an experience. However, wearers' seldom wish to review life events by browsing large collections of images manually (Ashbrook et al. 2006, Lin & Hauptmann 2006, Bell & Gemmell 2007, Lee & Dey 2007). The challenge tackled in this thesis then is to *manage*, *organise* and *analyse* these large image collections in order to interpret and extract knowledge for practical use of lifeloggers and others.

In this thesis, we argue that the lifelogging process is a Complex System as are the data collected (Li, Crane & Ruskin 2013, Li, Crane, Ruskin & Gurrin 2013*b*, Li et al. 2014, Li, Crane, Ruskin & Gurrin 2013*a*, Li, Crane, Gurrin & Ruskin 2016, Li, Gurrin, Crane & Ruskin 2016). After all, Complex Systems are often affected by a variety of phenomena at multiple temporal scales and the data, which typically display non-linear dynamics, are unpredictable, non-stationary and high dimensional. From the Internet to the global ecosystem, from road traffic networks to stock markets, from biological to social systems, time series methods are often employed to monitor and analyze complex, dynamic systems, which exhibit atypical or non-stationary characteristics (Hamilton 1989, Gopikrishnan et al. 2000, Schütte & Huisinga 2003, Horenko et al. 2008, Putzig et al. 2010, Vespier et al. 2012).

To add difficulties of analysis, lifelogging data is also affected by external factors such as temperature, movement, light and so on. Despite this, these data record repeated patterns at different resolutions due to the device's response to recurring events such as everyday life activities. In consequence of these characteristic features, lifelogging is an ideal test-bed for evaluating the

methods we introduce, which typically have been applied previously to other domains, (as above).

1.2 Objectives

In this dissertation, time series methods have been used to identify and extract useful information from temporal lifelogging image data, for which we do not have the benefit of prior knowledge. Important aspects include: noise reduction, structure and characterization of the raw data; the detection of multiscale patterns; and the mining of important, previously unknown repeated patterns. To achieve this, we have developed the following set of specific sub-goals or objectives:

1. To detect presence of any strong long-range correlation relationships in image time series (through Detrended Fluctuation Analysis (DFA)).

2. To demonstrate atypical or non-stationary characteristics in image time series (through investigation of equal-time Cross-Correlation Matrix).

3. To address the issue of noise reduction in the Cross-Correlation Matrix (using Random Matrix Theory (RMT)).

4. To characterise the 'most important' or 'unusual' events to the wearer (through analysis of the associated dynamics of the eigenspectrum).

5. To identify feature patterns characteristic of an individual's 'lifestyle' (using a motif discovery technique in image time series).

6. To test our motif discovery technique (by application to two known lifelog data collections, in this case the AIHS and NTCIR-12 Lifelog).

1.3 Outline of Thesis

Here, we give a brief outline of the dissertation, summarizing the contents of the following chapters. As most chapters are based on previous publications by the author, we also give the appropriate references, where applicable.

In **Chapter 2**, (Background) we outline the research context, review previous work and introduce some of the ideas followed in the thesis.

In **Chapter 3**, (Multiscaled Cross-Correlation Dynamics) we first introduce the Detrended Fluctuation Analysis (DFA) method, which aims to detect *long-range correlation* relationships in image time series. Equal-time Cross-Correlation Matrix is analysed to characterise dynamical changes in non-stationary multivariate image time series. Finally, the Maximum Overlap Discrete Wavelet Transform (MODWT) is used to calculate the equal-time Cross-Correlation Matrix over different time scales and to examine the details of the eigenvalue spectrum. This work has been published in the following papers:

- Li, N., Crane, M. & Ruskin, H. J. (2013), 'Automatically Detecting "Significant Events" on SenseCam', *International Journal of Wavelets, Multiresolution and Information Processing* 11(06), 1350050.

- Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2013b), 'Multiscaled Cross-Correlation Dynamics on SenseCam Lifelogged Images', *International Conference on Multimedia Modeling*, pp. 490–501.

In **Chapter 4**, we consider use of Random Matrix Theory (RMT) in order to investigate noise reduction in the Cross-Correlation Matrix. This work has been published in the following papers:

- Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2013a), 'Application of Statistical Physics for the Identification of Important Events in Visual Lifelogs', *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 589–592.

- Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2014), 'Random Matrix Ensembles of Time Correlation Matrices to Analyze Visual Lifelogs', *International Conference on Multimedia Modeling*, pp. 400–411.

In **Chapter 5**, (Finding Motifs in Large Personal Lifelogs) we introduce a motif discovery technique to explore detection of recurring and recognizable episodes of an individual's image data. These discovery techniques are applied to two lifelog data sets, the AIHS and NTCIR-12 Lifelog, in order to examine multivariate recurrent patterns of multiple-lifelogging users. This work has been published in the following papers:

- Li, N., Crane, M., Gurrin, C. & Ruskin, H. J. (2016), 'Finding Motifs in Large Personal Lifelogs', *Proceedings of the 7th Augmented Human International Conference*, pp. 9:1–9:8.

- Li, N., Gurrin, C., Crane, M. & Ruskin, H. J. (2016), 'NTCIR-12 Lifelog Data Analytics', *Proceedings of the first Workshop on Lifelogging Tools and Applications*, pp. 27–36.

In **Chapter 6**, Conclusions and Future Work, we provide a summary of the research to date, our conclusions and a discussion of potential future work.

Chapter 2

Background

2.1 Introduction

In this chapter, we elaborate on the history of lifelogging and its applications and review prior work on visual lifelogging, followed by a survey of event segmentation literature. Introducing the scientific field of Complex Systems, we describe some of the common characteristics found in nature, society and science. The features that are common to visual lifelogging and its representation as a complex system, such as high correlation, non-linear dynamics and multiple time scales, are briefly described. Next, we review several time series concepts and methods that are proposed for organising, analysing and managing lifelogging data sets. Finally, we introduce the data sets that are used in this thesis to test our methods.

2.2 History and Applications of Lifelogging

Lifelogging has a very long history: it can after all, arguably be traced back to 1917, when Fuller created a very large scrapbook called "Dymaxion Chronofile" (Harwood 2009), which he subsequently used to document all correspondence, manuscripts, drawings and audio-visual material and other documents relevant to his personal and professional life. This record of his personal archive divided into 15-minute increments is now the center piece of the Buckminster Fuller Archive at Stanford University.

The concept of logging and storing all of an individual's accumulated data (in digital format) was first proposed by the American engineer Bush in 1945 under the article "As We May Think". He described a hypothetical computer system called "Memex", as illustrated in Fig. 2.1. The Memex would provide an "enlarged intimate supplement to one's memory". Bush describes a Memex as a desk-based mechanical device that would allow individuals, specifically researchers, to develop their personal research archives by creating and following associative links or trails, allowing content to be used for their own individual use or for sharing with others (Bush & Think 1945, Buckland 1992). Although this device was never built, the concept of the Memex influenced the development of early personal knowledge base software.

Early research into lifelogs started in the 1980s when Mann began using wearable computing devices to create a record of his life. He performed significant studies on visual memory prosthetics that transformed his eye into a camera and his body into a Web server (Mann 1997, 2004). He refers to this technique as "cyborg logging" or "glogging"[2]. As one of the pioneers in the

[2]http://wearcam.org/glogs.htm

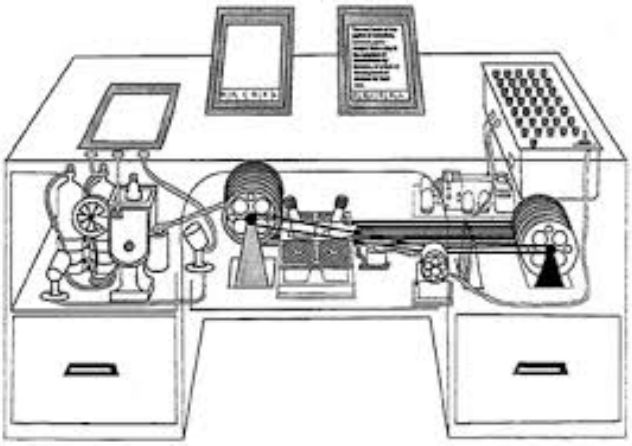

Figure 2.1: The **Memex**. Source: *Bush & Think (1945)*

field in lifelogging his research has led to the development of novel types of sensors and display hardware.

Another milestone in lifelogging research was the "MyLifeBits" project sponsored by Microsoft in 1998 to explore the potential of digitally chronicling a person's life (Gemmell et al. 2002, 2006). MyLifeBits was inspired by the above-mentioned Memex idea. The project focuses on preserving the life-profile of veteran researcher Bell. Using an infrared "SenseCam" camera worn around his neck, as well as scanners, and computing devices, Bell has captured a lifetime's worth of photographs, pictures, presentations, home movies, videotapes lectures, voice recordings, articles, books CDs, letters, memos and papers into the MyLifeBits software (Bell & Gemmell 2007). The MyLifeBits software supported full-text search, text and audio annotations, hyperlinking, reporting, visualising and clustering between content. This project shows the origins of the idea of lifelogging and digital self-tracking. The goal technologically is to create a personal archive, or a "portable, infallible, artificial memory" that can be exploited to increase job productivity, serve as a basis

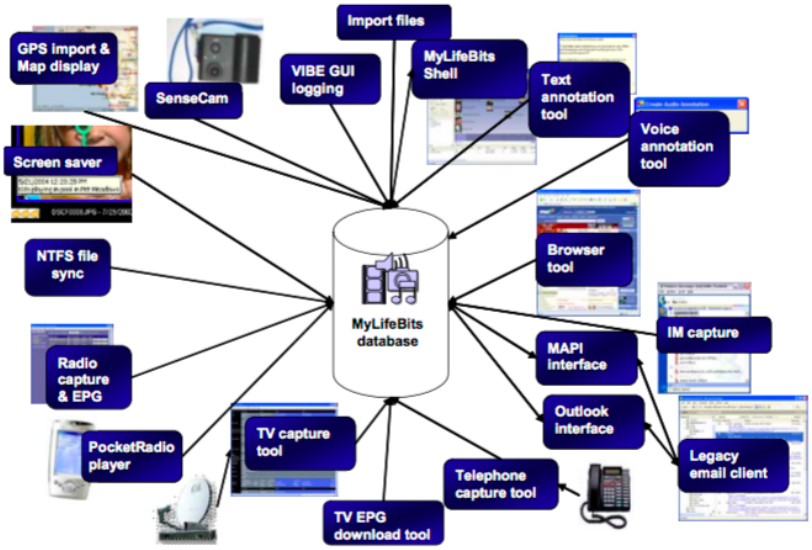

Figure 2.2: Overview of MyLifeBits. Source: *Gemmell et al. (2006)*

for medical treatment or to improve performance in school and in many other scenarios.

Much of the early research into lifelogging has focused, therefore, on developing technology to automatically populate a computer-based storage of life experiences in as much detail as possible. Until around 2007 there were no commercial lifelogging apps available but, through the emerging ubiquity of smartphone platforms (notably iOS and Android) and their associated app stores, we now have hundreds of lifestyle apps, many in the health and fitness sector with a significant number focussed specifically on lifelogging. While sensors such as heart rate monitors and pedometers have been around for more than ten years, the automated recording and especially processing of such data into a lifelogging format is more recent (Li & Hopfgartner 2016).

The technology research and advisory company Gartner predicts that, by 2020, the analysis of consumer data, collected from wearable devices, will be

the foundation for up to five percent of sales from the Global 1000 companies (Franklin 2013). Given these predictions, it comes as no surprise that more and more companies such as Apple, Fitbit, Garmin and Huawei entered the market with novel wearable devices. Due to improved sensing devices, increasing processor power, reducing cost of data storage and improvements in speed of data networks and battery technology, a multitude of devices, services and apps are now available that track almost everything we do. Lifelogging apps of note include *Saga, Track my Life, Instant, Journaly, Loca, Fit Time and Sleeply.* In addition to the phone apps there has been a proliferation of wearable technologies such as Fitbit[3], Google Glass[4], Apple Watch[5] and many more that can track activities and specific aspects of a subject's life.

2.3 State-of-the-Art in Visual Lifelogging

Wearable cameras play an important role in research on lifelogging since analysing camera data streams can reveal a lot of information. *Visual Lifelogging* is a form of lifelogging based on the passive capture of photographs of a person's experience. Fig. 2.3 shows some pictures taken by author-wearable camera, the SenseCam device, over the course of a day.

[3]https://www.fitbit.com/ie/
[4]https://www.x.company/glass/
[5]https://www.apple.com/ie/watch/

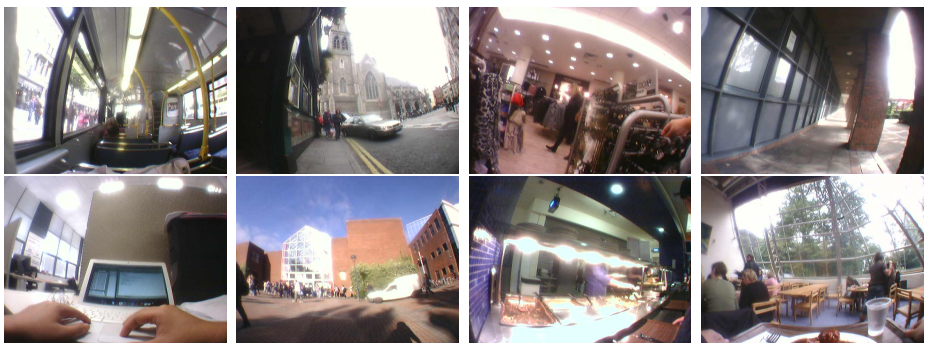

Figure 2.3: Examples of images taken by author-worn SenseCam.

Visual lifelogging as such also dates back to early work of Mann, who led the Digital Eye project in the 1980s that aimed to develop the first customised lifelogging and ubiquitous computing hardware, called *EyeTap* (see Fig. 2.4 (a)). This device allows "the user's eye to operate as both a monitor and a camera as the *EyeTap* intakes the world around it and augments the image the user sees." Mann has gone on to develop many generations of wearable camera devices for recording aspects of his life by means of continuously improving technology. In 1994 Mann began to record his life 24 hours per a day and 7 days per week. The *EyeTap* addressed many of the fundamental challenges found in wearable lifelogging (personal imaging (Mann 2004)). From the observer's perspective, *EyeTap* looks quite similar to today's Google Glass as shown Fig. 2.4 (b).

Visual lifelogging gained in popularity when Microsoft introduced their SenseCam (as shown in Fig. 2.4 (c)) to the market in 2006. Motivated by an accident causing memory-loss, the designer wished to reinforce recall by use of an automatic device that could automatically record and save everything that occurred (Selke 2016). The SenseCam is a small and lightweight device that is designed to be worn on the chest and to take images and sensor readings

passively, without user-intervention. Normally, the SenseCam takes a picture every 30 seconds and thus thousands of pictures will be collected per day. Wearable Cameras such as Microsoft SenseCam and the more recent Narrative Clip (as shown in Fig. 2.4 (d)), are producing low temporal resolution (LTR) images (e.g. 2-4 frames per minute) that allow capturing of images over a long time period without the need to recharge the battery. Consequently, they are a promising source for inferring details such as behaviour patterns, habits or lifestyles of the user. Other devices like the GoPro and Google Glass are wearable cameras that are commonly mounted on the head. These devices have a relatively high temporal resolution (HTR) (e.g. 25-60 frames per second); in other words, they can capture smooth continuous video during certain moments of time with high quality resolution. Consequently, these devices offer potential for in-depth analysis of daily or special activities.

The visual analysis of the HTR wearable cameras (continuous video) data can be treated as a conventional videos processing task. The recent development of deep learning techniques in the computer vision field has resulted in powerful approaches to extract meaningful context from video material. However, lifelogging cameras with LTR are generating temporally ordered images taken over the course of a day, not continuous video. Therefore, we cannot just simply adopt deep learning techniques for analysing these temporal images data sets. They are much more challenging than the conventional videos due to the nature of the data sets: free motion of the camera with low resolution images, long and unstructured images sets, motion blur and rapid illumination changes (very noisy), large number of non-informative images, etc (Tan et al. 2014). Hence, computer vision techniques based on feature extraction have

(a) EyeTap (b) Google Glass (c) SenseCam (d) NarrativeClip

Figure 2.4: A range of the popular wearable cameras. Source: *Streams (2012), Michael et al. (2018)*

proven to be rather limited for such photographic cameras data streams.

In this thesis, we focus on *Visual Lifelogging* created by low temporal resolution cameras such as SenseCam. The SenseCam has received a lot of attention from scientific researchers with more than three thousand papers on various aspects of lifelogging since 2005. It has been shown that such lifelogging images and other data can be periodically reviewed to recall and strengthen an individual's memory (Hodges et al. 2006, Sellen et al. 2007, Berry et al. 2007, Arcega et al. 2013). Lifelogging enables memory to be searchable, retrievable and shareable. The use of wearable cameras such as SenseCam has been explored, so far mainly for memory enhancement, but also for health monitoring and well-being, in social iterations and for leisure activity records amongst others (Fleck & Fitzpatrick 2009, Aghaei et al. 2018). A recent study suggests that lifelogging has even wider application such as detecting stress in real-world driving (Dobbins & Fairclough 2018), analysing human nutritional habits (Aguilar et al. 2018) and that an increasing number of devices are likely to be available in the near future (Askoxylakis et al. 2011).

One observation that can be made is that wearable cameras such as SenseCam generate a very large amount of data for a single typical lifelogger's day. Since many will contain similar features, not all pictures are needed to capture

moments of the lifelogger's life. The challenge with lifelogging, especially visual lifelogging, is extracting meaningful insights for the users. To date, much of the research on lifelogging has focused on aspects such as developing sensors, capture and storage of data (Mann et al. 2005, De Jager et al. 2011, Qiu et al. 2012), processing data into annotated events (Lin & Hauptmann 2006, Doherty & Smeaton 2008, Doherty et al. 2008, Spriggs et al. 2009, Li, Crane & Ruskin 2013, Lu & Grauman 2013, Bolaños et al. 2014, Poleg et al. 2014, Talavera et al. 2015, Castro et al. 2015), search and retrieval of information (Wang et al. 2006, Aghazadeh et al. 2011, Wang & Smeaton 2012, Chandrasekhar et al. 2014), assessing user experience and designing user interfaces for applications of the memory aids (Hodges et al. 2006, Sellen et al. 2007, Berry et al. 2007, Arcega et al. 2013), diet monitoring (Reddy et al. 2007), analysis of activities of daily living (ADL) (Danna & Griffin 1999) and so on. However, although some of these challenges have been comparatively well addressed (Hodges et al. 2006), resulting in improved wear-ability of devices and inexpensive storage, (Gemmell et al. 2006), the challenge has now shifted to *interpretation* and *extraction* of knowledge from the vast quantities of captured data. Research on visual lifelogging data mainly focuses on five areas: Data Acquisition, Informative Image Detection, Temporal Segmentation, Egocentric Summarization and Content-Based Search and Retrieval (Bolaños et al. 2017). In this thesis, we are principally concerned with Temporal Segmentation, in an effect to render tractable the large volume of lifelogging data.

Motived by Lin and Hauptmann's argument that "...continuous video needs to be segmented into manageable units ..." (Lin & Hauptmann 2006), segmenting visual lifelogging data into different events or activities offers an ideal way

to organise and structure lifelogging images. Previous work has been carried out on the segmentation of lifelogs by Doherty et al. (2007) who used a wearable camera's readings such as low-level image descriptors, audio, temperature, light and movement data to segment the SenseCam images into distinct activities. However, the accuracy of automatic detection varies greatly and an analysis of *specific* activities is still lacking. Later on, Doherty & Smeaton (2008) improved their previous work on event segmentation by incorporating time periods (morning, afternoon and evening) as crude event boundaries. They segmented SenseCam images into approximately twenty distinct events that occur in a wearer's day, which translates to over 7,000 events per year (as illustrated in Fig. 2.5). Nevertheless, this large collection of personal information still contains a significant percentage of routine events and an analysis of individual *specific* activities on natural activity boundaries is still lacking. More recent work carried out by Wang & Smeaton (2013) focused on the idea of developing an ontology framework by mapping low-level features like colour and texture, to high-level concepts such as 'indoor', 'outdoor', 'driving', 'eating' etc. Although this method improves accuracy, it is still important to note that the approach is based on training classifiers on a set of annotated *ground truth*[6] images. Subsequently, an automatic egocentric video segmentation was developed by applying the R-clustering segmentation approach based on a convolutional neural network (CNN, or ConvNet) technique to compute the whole image as a fixed feature extractor for image representation (Talavera et al. 2015). The authors offer the possibility to integrate different clustering and segmentation methods, giving more robust results. More recently,

[6]Ground truth is a term used in statistics and machine learning that means checking the results of machine learning for accuracy against the real world.

researchers have looked at the relationship between lifelog images and non-lifelog images, by applying a domain-adversarial CNN deep learning model to the domain of visual lifelogs in order to transfer knowledge from the domain of visual non-lifelog data. The accuracy of this approach depends heavily on the number of training examples (Ye 2018).

To date, advances in visual Lifelogging analysis is driven by computer vision research. However, the automatic visual analysis of lifelogs remains a significant challenge due to two main reasons: On the one hand, wearable cameras have a small field of view which produces low temporal resolution images; On the other hand, the free motion of wearable cameras result in most images often being rotated, blurred, or capturing fast illumination changes. Many computer vision researchers have tried to extract low-level features such as colour histograms, textures, and shapes from visual lifelogging images. Although the computational cost for extracting low level features are incredibly cheap, researchers struggle to achieve high accuracy for automatic visual analysis. More recently, deep learning networks have been heavily applied by computer vision researchers to classifying and tagging images and the level of accuracy has improved significantly. However, the effective transfer of learned models to different application domains such as visual lifelogs is still questionable (Wang et al. 2018). This is due to the quality of lifeloggging images, as outlined above.

All methods mentioned above apply machine learning techniques or deep learning techniques. The drawback of these techniques is that they require *prior* training of models, so that features which do not appear in the training data set tend to be less well detected in the new data. Given that visual lifelogs usually consist of large and often unstructured collections of multimedia

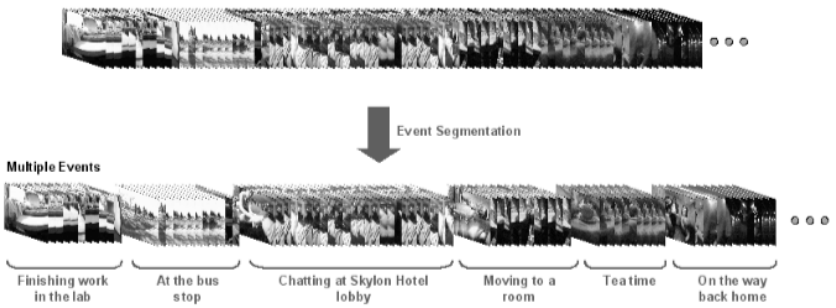

Figure 2.5: Segmenting a day of SenseCam images into distinct events. Source: *Doherty (2009)*

information, such rule-based and concept-based techniques may not offer the best solution for this content. Furthermore, for privacy reasons, the amount of visual data that can be used for training purposes is very limited as not many lifeloggers are willing to share their personal images. Even those who are willing to do so may partially share only or perform some pre-processing tasks to anonymise the data (Gurrin et al. 2016*a*). Another limitation is that all the training data has to be labelled[7], which is extremely costly due to the large volume of images. Moreover, due to the free motion of wearable cameras and passive acquisition, lifelogging data presents huge variability in terms of illumination conditions and object appearance, i.e. considerable *noise*. Given all these limitations, we argue that concept-based and rule-based methods for segmenting lifelog sequences into a set of discrete events are not suitable for all use cases. In this thesis, we address this by introducing an approach to organise and manage such varied and large data sets without prior knowledge.

[7]In machine learning, labeled data is a group of samples that have been tagged with one or more labels.

2.4 Complex Systems

The field of Complex systems is an interdisciplinary one concerned with description of systems composed of many interacting components that interact with each other across many different temporal and spatial scales. The concept has played an important role in studying natural phenomena, physical and artificial or hybrid systems. Application examples from different domains include the immune system (Varela & Coutinho 1991), the human brain (Rubinov & Sporns 2010), flocking or schooling behaviour in birds or fish (Sumpter 2010), road traffic flows (Lämmer & Helbing 2008), the economy and financial markets (Sornette 2017), cloud platform based on online services (Patel et al. 2013) and complex software based on large-scale and distributed architectures (Lou et al. 2013). However, there is no precise definition of a Complex System. We attempt to characterize them here and include a core set of features that are widely associated with them (Kantelhardt 2009, Newman 2011, Ladyman et al. 2013):

- *Agent-Based*

  The system is composed of independent agents which assess their situation and make decisions on the basis of a set of rules.

- *Nonlinearity Dynamics*

  The system behaves in an unpredictable and chaotic manner due to agents responding to their environment in different ways. For this reason the system displays highly non-linear dynamics and is why it can not just simply be described by linear equations.

- *Emergence*

The appearance of patterns, structures or properties occurs due to co-evolutionary behaviour or self-organisation; emergence cannot be previously observed or planned from functional characteristics of the system. This implies that direct forecasting is impossible. Uncertainty analysis and sensitivity analysis are commonly used approaches for evaluating stochastic outcomes.

- *Feedback*

  The process whereby a system variable influences another variable, either positively or negatively.

- *Fractal (Multifractal) Structure*

  A pattern or a structure that is self-similar at different spatial scales.

- *Critical Points*

  The accumulation of small events can lead to main changes. Once the aggregation of small stimuli reaches a particular threshold *(the critical point)* large fluctuations in the system can ensue.

- *Adaptation*

  The agents have the ability to learn from and adapt to each other. They are able to change their behaviour under different circumstances and evolving.

In this thesis, lifelogging data (which display a number of the characteristics above, such as high correlation, uncertainty, non-stationary, nonlinearity and diverse behaviour at different temporal scales) are, to our knowledge, considered as a complex system for the first time. However, the understanding and characterisation of the complex system is a challenging task, since splitting

a system into simpler subsystems or single components is not possible without losing dynamical properties. One approach in studying such systems is the recording of long time series of several selected observed variables, which reflect the state of the system in a dimensionally reduced representation (Kantelhardt 2009, Meyers 2011). In this thesis, therefore, the analysis of complex image sensor data in the form of time series is of primary interest. Usually, the data generated by complex systems exhibit fluctuations or non-stationarity over a wide range of time scales. In order to observe the data's dynamical properties, and draw on earlier research in differ domains, several methods originating from Statistical Physics have been used for studying the statistical characteristics and examining the structure and dynamics of lifelogging images time series. The goal of these methods is to identify and extract useful knowledge from these data.

2.5 Time Series Analysis

Given the recent rapid development of data collection and data storage, we can easily collect vast quantities of sensory temporal data with large dimensionality, (commonly used to detect the dynamic change of some specific Complex System). Time series data are recorded in many areas of science and engineering including signal processing for finance, biosystems and chemical processes, and statistical analysis of such series is widely used for extracting characteristics of data sequences (Ding et al. 2008). Before introducing the approaches for analysis of lifelogging images, we first review some relevant time series concepts. An understanding of these concepts provides motivation for the methods introduced. Results are discussed in the remainder of the chapter.

2.5.1 Stationary and Non-Stationary Time Series

Stationary time series are those for which statistical properties such as mean signal value, variance, or autocorrelation are constant over time (Adhikari & Agrawal 2013). A given series thus remains in relative equilibrium in relation to its corresponding mean value. In contrast, non-stationary is exhibited when statistical properties change over time. A majority of applications create non-stationary time series. Such applications include financial markets (Hamilton 1989, Putzig et al. 2010), biophysical systems (Schütte & Huisinga 2003), climate systems (Horenko et al. 2008), and so on.

2.5.2 Univariate and Multivariate Time Series

A time series is described as a set of data points, measured typically over successive time points, where time implies a recognizable continuum. It is mathematically defined as $X_i(t)$ where $[i = 1, .., n; t = 1, .., m]$, where i indexes the measurements made at each time point t (Cochrane 2005, Box et al. 2015). A univariate time series has n equal to unity, and a multivariate time series has n equal to, or greater than, two. A multivariate time series item is typically stored in an m x n matrix, where m is the number of observations and n is the number of variables. Generally speaking, a univariate time series has only one variable that varies over time, while multivariate time series have several. Multivariate time series data sets are commonly found in various domains such as finance (Laloux et al. 1999, Plerou et al. 1999, 2000, Siwy et al. 2002, Jánosi & Müller 2005, Santhanam et al. 2006, Carpena et al. 2007), electroencephalography (EEG) (Gopikrishnan et al. 2000), magnetoencephalography (MEG) (Utsugi et al. 2004). Modelling of a univariate time

series is well-established, but, in practice may be of limited value as most experiments collect a multiplicity of data sets. The most commonly used method for the modelling and interpretation of multivariate time series is the study of cross-correlation relationships among the time series. For example, in finance, cross-correlation among different financial markets has been studied to improve the predictability of financial return series (Duan & Stanley 2011) and for risk management (Conlon et al. 2009). In this dissertation, for the lifelogging device, we consider the generated image sequence as a time series, which exhibits non-stationary fluctuations across a wide range. Although visual lifelogging devices like the SenseCam normally take a picture every 30 seconds, such images are also affected by different external temporal factors at different scales such as light, temperature and so on. Moreover, the lifelogging image data set displays repeated patterns at different resolutions due to recording individual local life activity and interactions at larger scales. Time series methods that are often used to structure and characterise non-stationary multivariate time series are discussed in what follows.

2.5.3 Detrended Fluctuation Analysis

Detrended Fluctuation Analysis (DFA), an example of a nonlinear signal analysis technique, has been used commonly to characterise power-law scaling in time series. It was originally developed for the evaluation of DNA sequences (Peng et al. 1993, 1994, Buldyrev et al. 1993, Stanley et al. 1999). However, in the last few decades it has been widely applied to many natural time series, generated by complex systems, including those in meteorology, (Koscielny-Bunde et al. 1998, Talkner & Weber 2000, Ausloos & Ivanova 2001), cardiac

dynamics (Ivanov et al. 1996, 1999), astrophysics (Moret et al. 2003), climate change (Pelletier 1997, Zheng et al. 2018), stock prices (Vandewalle & Ausloos 1998, Gopikrishnan et al. 1999, Ausloos & Ivanova 2000, Caraiani 2012) and many others. More recently, DFA has also been applied to detect activity fluctuations at different time scales created by accelerometer sensors worn by elderly people (e.g. those with Alzheimer's Disease (Hu et al. 2014)). DFA has thus been established as an important tool for the detection of long range auto-correlations in time-series with non-stationary signals. The advantage of DFA is that it can both remove different trends from external factors in the data and reduce noise level measurement. The images captured from lifelogging devices are characterised by many examples of the former: multiple sensor data, an accelerometer to detect motion, sensors to detect changes in light levels and so on. Image time series are consequently complex and composed of many interacting units. The DFA method can thus be adapted initially to analyse these and remove stationary trends, helping to highlight non-stationary events, which could be of importance. The DFA method can be used, therefore, to provide valuable initial background analysis for lifelogging image time series.

2.5.4 Cross-Correlation Analysis and Noise Reduction

Cross-Correlation analysis, as noted previously, is in common use when dealing with multivariate time series. The behaviour of the largest eigenvalue of a cross-correlation matrix over small windows of time has been studied for financial series (Laloux et al. 1999, Plerou et al. 1999, 2000, Siwy et al. 2002, Jánosi & Müller 2005, Santhanam et al. 2006, Carpena et al. 2007), electroencephalographic (EEG) recordings (Gopikrishnan et al. 2000), mag-

netoencephalographic (MEG) recordings (Utsugi et al. 2004) and a variety of other multivariate data. Drawing on these applications in this thesis, we adopt a similar approach to investigate the dynamics of Lifelogging image sequences.

However, as available series are typically of limited length, using such finite series to estimate cross-correlation results in a correlation matrix containing much which corresponds to "random" contributions (Wigner 1951, Dyson 1962, Dyson & Mehta 1963, Mehta & Dyson 1963). This phenomenon can also be observed in other domains including number theory and combinatorics (Conrey et al. 2005), wireless communications (Tulino et al. 2004), and in multivariate statistical analysis and Principal Components Analysis (Ulfarsson & Solo 2008), as well as for financial and other large dimensional data analysis (Plerou et al. 2002, Conlon et al. 2007). A well-proven technique to handle the issue of noise and to help reduce this is the application of Random Matrix Theory (RMT) (Plerou et al. 2002).

Random Matrix Theory (RMT) was first introduced by Wigner, Dyson, Mehta, and others who aimed to study the energy levels of complex atomic nuclei (Plerou et al. 2002). RMT predictions represent an average over all possible interactions. Deviations from the *universal* predictions of RMT can be used to identify system-specific, non-random properties of the system under consideration, providing clues about the underlying interactions (Dyson 1962, Dyson & Mehta 1963, Mehta & Dyson 1963). Applications to analysis of the spectral properties of an empirical correlation matrix show that a large proportion of the eigenvalues of the cross-correlation matrix agree with RMT predictions, (indicating a considerable degree of randomness in measured cross-correlation), but major deviations are also observed.

2.5.5 Wavelet Analysis

An additional mathematical tool used here is the wavelet transform (WT), that can decompose time series into different frequency components (sub-series), enabling the study of each component with a resolution matched to its scale. This method has been applied to studying the dynamical features of non-stationary time series and has served as the most important signal processing tool in image analysis (Zhu 2008), meteorology (Can et al. 2005), and financial time series (Bouchaud & Potters 2003). Specifically, it has been used to decompose a signal into different time scales (time components) instead of frequencies, as is the case with the Fourier transform. As it is composed of basis functions that are discontinuous, it can be used to analyse time series which contain discontinuities and sharp spikes. Further, the wavelet technique can be used to zoom in on detailed features of the time series and to construct the whole picture of a time series at different 'window' sizes [8]. The wavelet transform focuses on two parameters, frequency (scale) and time. Fig. 2.6 shows the time-frequency resolution property of the wavelet transform. It demonstrates that a narrow window variable yields information on high frequency movements, while a wide window yields information on low frequency movements.

Wavelet methods have achieved an impressive popularity for studying interdependence of economic and financial time series and have been widely used since such data contain layered structure, each feature occurring at a different time scale. In applying wavelet methods to a multiresolution analysis of the high frequency stock index, it has been shown that a wavelet-matching pursuit

[8] Long time intervals when a more precise low frequency information is needed, and shorter intervals when high frequency is needed.

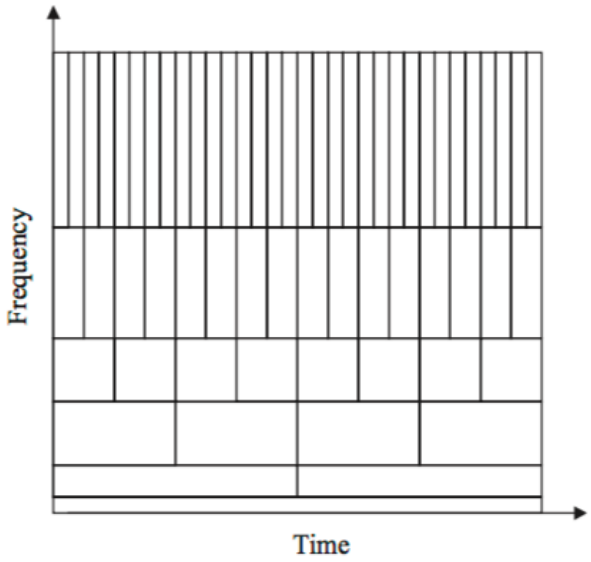

Figure 2.6: The wavelet transform partitioning of frequency and time plane. Source: *Gallegati (2008)*

algorithm can be used to uncover hidden periodic components (Capobianco 2004). For the investigation of scaling properties of foreign exchange rates using Maximum Overlap Discrete Wavelet Transform (MODWT - a method for the modification of ordinary discrete wavelet transform), Gençay et al. (2001), found for example, that foreign exchange rate volatilities can be described by different scaling laws over different horizons. Similarly, wavelet-multiscale studies have also been reported (Conlon et al. 2008) where the authors have examined Wavelet multiscaling for medium and high-frequency intra-day stock returns. In this thesis, we apply a similar technique to investigation of the atypical or non-stationary characteristics in lifelogging image time series at different scales, in order to highlight the 'most important' or 'unusual' events for the lifelogger.

2.5.6 Time Series Motifs

Taking this idea of sub-patterns further, many researchers have worked recently on the extraction of various repeated short - to mid-length characteristics of time series. Keogh & Lin (2005) have demonstrated that finding such time series *motifs* offers a promising solution for extracting meaningful results in a massive time series' database. Previously unknown or seldom applied in a general context, motif analysis seeks to identify frequently occurring patterns in time series data, where these have proved powerful for various mining tasks including clustering, classification, summarization and visualization (Lin et al. 2002, Chiu et al. 2003). Fig. 2.7 illustrates an example of a motif discovered in an astronomical database, showing repeated patterns.

Time series motifs, stemming originally from genomic analysis ideas, such as selection of maximally informative genes (Androulakis 2005), protein sequence identification (Nevill-Manning et al. 1998) have been widely applied also in domains as diverse as measuring data from body sensors (Minnen et al. 2006), finding patterns in sports in motion capture data (Tanaka et al. 2005) as well as in video surveillance applications (Hammid et al. 2005).

Many researchers have proposed algorithms for identifying time series motifs (Lin et al. 2002, Chiu et al. 2003). SAX (Symbolic Aggregate approXimation); (Lin et al. 2003), in particular, has been extremely successful due its simplicity and efficiency for representation. It has been readily adopted in various domains such as bioinformatics, information retrieval and language processing (Lin et al. 2003, Ross et al. 2012, Lagun et al. 2014). The SAX technique transforms a numerical time series into a sequence of symbols representing values by a finite alphabet. This method is very simple and does

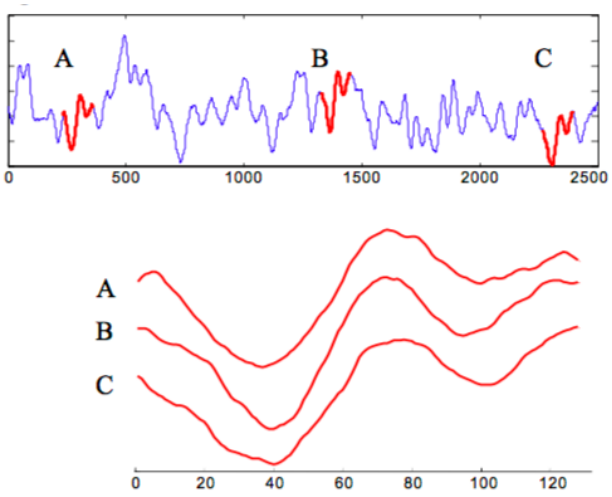

Figure 2.7: An astronomical time series contains 3 near identical subsequences. Source: *Lin et al. (2002)*

not require any *a priori* information about the input time series. Fig. 2.8 illustrates ideas of the SAX method, where the algorithm consists of three steps:

- Divide a time series into segments of length L
- Compute the average of the time series for each segment
- Quantize the average values into a symbol from an alphabet of size N

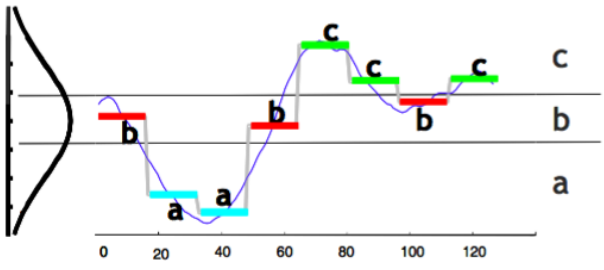

Figure 2.8: The SAX motif discovery technique. The different letters represent repeated motifs or patterns. Source: *Lin et al. (2007)*

SAX achieves dimensionality reduction and indexing with a lower-bound distance measure, providing fast answers whether exact or approximate. However, it is less suitable for extraction of characteristic features from multi-dimensional time series data. Consequently, Principal Component Analysis (PCA) has been used to transform multi-dimensional time series into a one dimensional realisation to detect motifs that are common to all (Tanaka et al. 2005). More recently, an *extended* motif discovery method has been reported for a single time series designed to detect motifs occurring across *several dimensions* of a multi dimensional signal (Minnen et al. 2007). Visual lifelogs contain records of a wearer's life activities that occur over different time periods and, in consequence, we argue that discriminating motifs can help represent activities of different length and timing.

2.6 Data Sets Used in this Thesis

Although the concept of lifelogging is not novel, the number of quality and quantitative activities is growing rapidly. This is mainly the result of a rising profile of lifelogging and its key aspects of being perceived as self-aware and personal archives. Given the nature of lifelogging data i.e., personal information content, privacy, security and ethical issues have been carefully concerned by lifeloggers as lifelogging images might capture sensitive information of the lifeloggers or reveal the identity of others who are captured by the wearable cameras (Li & Hopfgartner 2016). To the best of our knowledge, prior to the release of the NTCIR-12 Lifelog data set, only two public lifelogging data sets consisting of Images were made available to the public, namely the "All I have Seen" (AIHS) (Jojic et al. 2010) and "The Egocentric Dataset of the

University of Barcelona" (EDUB) (Bolaños et al. 2018) data sets. It is noted that most lifelogging image data collections are far from perfect, due to incompatible Lifelogging devices, errors occurring during the lifelogging process and passive acquisition of lifelog material. Therefore, the construction of rich data sets and accurate annotations become crucial for the development of analytic methods on visual lifelogging (photographic cameras).

Four data sets have been used for our study. The first data set (D_1) (D_1= AUTHOR6DAYS) was generated by the author, who wore the SenseCam over a six day period, from a Saturday to a Thursday, forming a total Lifelog of 10,260 images, the average wearing time varied from about 6 hours to 11 hours between weekends and weekdays respectively. Descriptive statistics are reported in Table 2.1. These particular days were chosen in order to include a weekend, where normal home activity varied in comparison to events on weekdays or within the (extended day of a) working week:

- Weekend

  Saturday, a typical example to illustrate the difference mentioned above, involved the subject walking to the nearest bus stop from home, a bus journey to the city centre, walking through local streets as well as a visit to a shopping centre. This day also involved dinner with a friend and a bus journey back to the original bus stop. On the next day, the subject only wore the SenseCam during the afternoon to the office: thus images described the journey of the subject from the home to office, a period spent working in front of the laptop and the return journey back home.

- Weekdays

  Over the next four days, these images described a typical day for the

subject: sitting in the office, talking with a colleague and sharing lunch in the cafeteria and so on. On some days, the subject wore the SenseCam home, so that it recorded the wearer's journey from the office to home, and the next morning from home to the office.

Table 2.1: The AUTHOR6DAYS data set (D_1)

User	Events Catalogue	Images
1	Working	6146
1	Walking Outside	1494
1	Shopping	826
1	Eating	658
1	Taking Bus	297
1	Others	839
		Total: 10,260

The second data set (D_2) (D_2=AUTHORTYPICAL) consists of 2,096 lifelog images, recorded by the author using a SenseCam over a 'typical day' of her life. It contains activities such as commuting to the office in the morning, sitting and working in the office at a desk, talking with colleagues and sharing lunch in the cafeteria, as well as commuting back home in the evening and so on. Descriptive statistics are given in Table 2.2.

Table 2.2: The AUTHORTYPICAL data set (D_2)

Event Number	Event Series	Number of Images
1	Travelling to work	38
2	Arriving in the office	21
3	Working	136
4	Chatting with people	107
5	Working	157
6	Walking in a building	29
7	Working	234
8	Going to the bank	108
9	Working	412
10	Lunch	148
11	Working	668
12	Leaving the office	38
		Total: 2096

The third data set (D_3) (D_3=AIHSWEDS) was obtained from a different wearer who wore the SenseCam continuously over an extended period capturing, on average, an image every 20 seconds. The data set is referred to as "All I have Seen" (AIHS) (Jojic et al. 2010). On some days, data were captured for a few hours only. In order to investigate the predominantly 'active' lifestyle of the wearer more accurately, these incomplete 'days' have been removed from our study. AIHS data set thus included 18 days with a total of 34,846 images. Descriptive statistics are reported in Table 2.3. The data set contains a mix of indoor and outdoor scenes. In order to investigate the wearer's 'typical day' lifestyle, we studied data recorded on four Wednesdays. Selecting data recorded in the middle of the week was deliberate, given the likelihood of a more regular routine. The third data set (D_3) (D_3=AIHSWEDS) selected subset consists of 7,549 images. Descriptive statistics are reported in Table 2.4.

Table 2.3: AIHS Data Set Overall Summary

User	Day Number	Date	Week Day	Images
1	1	21-04-2009	Tuesday	1656
1	2	22-04-2009	Wednesday	1913
1	3	23-04-2009	Thursday	3286
1	4	24-04-2009	Friday	1619
1	5	28-04-2009	Tuesday	1683
1	6	29-04-2009	Wednesday	1921
1	7	30-04-2009	Thursday	1726
1	8	02-05-2009	Saturday	2100
1	9	03-05-2009	Sunday	1517
1	10	06-05-2009	Wednesday	2095
1	11	07-05-2009	Thursday	2019
1	12	09-05-2009	Saturday	2842
1	13	11-05-2009	Monday	1284
1	14	13-05-2009	Wednesday	1654
1	15	14-05-2009	Thursday	1097
1	16	17-05-2009	Sunday	2267
1	17	18-05-2009	Monday	1433
1	18	13-06-2009	Saturday	2734
				Total: 34,846

Table 2.4: The AIHSWEDS data set (D_3)

User	Date	Week Day	Images
1	22-04-2009	Wednesday	1913
1	29-04-2009	Wednesday	1921
1	06-05-2009	Wednesday	2095
1	13-05-2009	Wednesday	1620
			Total: 7549

The fourth data set is from the public domain, namely the *NII Testbeds and Community for Information access Research (NTCIR-12)* Lifelog data set[9]. The NTCIR-12 Lifelog data set was initially created for supporting the Information Retrieval (IR) community to develop new and novel lifelogging retrieval and visualisation systems, i.e. as a test baseline. The vast majority of published Information Retrieval (IR) research assesses effectiveness using resources known as *test collections*, in conjunction with evaluation measures. Test collections have a history dating from the 1990s: *Text REtrieval Conference (TREC)*[10] in US, *Cross-Language Education and Function (CLEF)*[11] in Europe and NTCIR in Asia run evaluation campaigns aimed at supporting the development, testing and evaluation of IR systems. In 2016, for the first time, NTCIR included The NTCIR-12 Lifelog as a pilot task (Kishida & Kato 2016). This represents the first data set test collection in lifelogging for the Information Retrieval community and the purpose of it as a pilot task was to explore methods of searching through large lifelog archives. The two

[9]http://ntcir-lifelog.computing.dcu.ie
[10]https://trec.nist.gov
[11]http://www.clef-campaign.org

subtasks are: 1) the Lifelog Semantic Access Task (LSAT); to examine search and retrieval from lifelogs and 2) the Lifelog Insight Task (LIT); to investigate knowledge mining and visualisation of lifelogs (Gurrin et al. 2016*b*).

NTCIR-12 Lifelog data were generated by three individuals, wearing the Autographer (as shown in Fig. 2.9)[12] camera for periods of about one month, capturing on average 1000-1500 images per day, where a sample image is also shown (Fig. 2.9). This camera uses five built-in sensors which include on accelerometer, magnetometer, temperature, color, PIR (infrared motion detector), and GPS. With a 5 megapixel low light image sensor and offering a 136 degree wide-angle lens, it captures images and other sensor readings automatically, recording the wearer's every moment. In the NTCIR-12 Lifelog, every image was resized down to 1024 x 768 resolution and all faces were blurred[13] manually. For our study, we selected a total of 34,758 images to include three lifelogging profiles over ten days from NTCIR-12 Lifelog data set to test our methods. Table 2.5 reports the full details of the fourth data set (D_4), where D_4=NTCIR-12 (3 USERS).

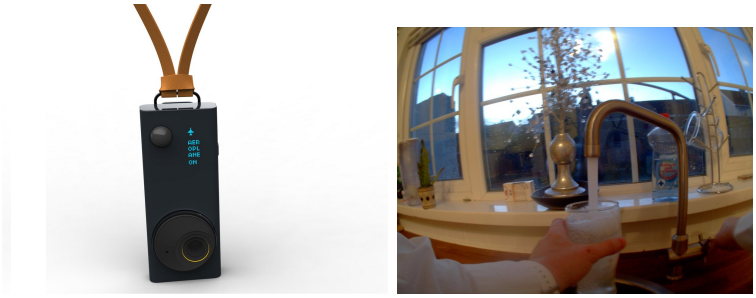

Figure 2.9: Autographer Wearable Camera and Sample; an image from the NTCIR12-Lifelog Test Collection

[12]https://en.wikipedia.org/wiki/Autographer

[13]Photo sharing can cause privacy issues. Face blurring is one method to deal with this issue.

Table 2.5: The NTCIR-12 (3 USERS) data set (D_4)

User	Date	Images	User	Date	Images	User	Date	Images
1	23-02-2015	1,392	2	20-04-2015	1,402	3	15-06-2015	1,130
1	24-02-2015	1,506	2	21-04-2015	1,038	3	16-06-2015	761
1	25-02-2015	1,570	2	22-04-2015	1,211	3	18-06-2015	985
1	26-02-2015	1,454	2	23-04-2015	1,100	3	19-06-2015	192
1	27-02-2015	1,280	2	24-04-2015	948	3	20-06-2015	460
1	28-02-2015	1,375	2	25-04-2015	1,004	3	21-06-2015	1,094
1	01-03-2015	1,527	2	26-04-2015	1,034	3	24-06-2015	1,299
1	02-03-2015	1,423	2	27-04-2015	1,389	3	28-06-2015	1,066
1	03-03-2015	1,129	2	28-04-2015	864	3	01-07-2015	956
1	04-03-2015	1,450	2	30-04-2015	1,089	3	11-07-2015	1,690
	Total:	14,106		Total:	11,079		Total:	9,573

2.7 Summary

In Chapter 2, we outlined the background to our research, introducing some of the difficulties inherent in analysis and the methods used to address these. We introduced the several data sets used for our research and described the sources. The history of lifelogging is extensively reviewed and seminal prior work on visual lifelogging - (in particular on event segmentation) is described. In this chapter, we also gave a broad introduction to features of Complex Systems and covered some relevant concepts and methods from time series and their analysis of similar features that arise in visual lifelogging. We first employ the Detrended Fluctuation Analysis (DFA) method to try to detect the statistical features in image time series. We then consider use of the Equal-time Cross-Correlation Matrix method to examine and characterise the structure and dynamic features of the data. We propose, further, the application of the Maximum Overlap Discrete Wavelet Transform (MODWT) technique to investigate the scaling properties of the series and the lead/lag relationships among the different time scales. The Random Matrix Theory (RMT) method also is applied to address noise reduction in the Cross Correlation Matrix. Finally, we introduce a novel motif discovery technique to extract patterns or motifs from image data. Fig. 2.10 illustrates the methodology covered by this thesis.

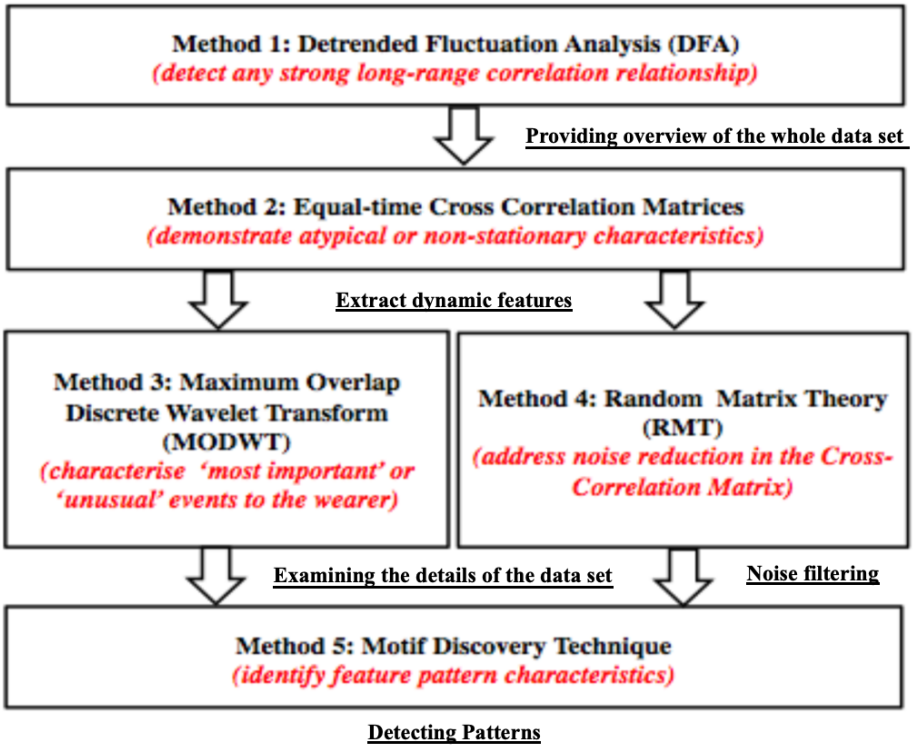

Figure 2.10: Outline of the Research Methods covered by this thesis

Chapter 3

Multiscaled Cross-Correlation Dynamics

3.1 Introduction

In Chapter 2, we reviewed methods that we proposed to use for analysis of lifelogging image data sets. The justification is that these methods are suitable as lifelogging captures primarily the sequence of a user's everyday activities, taking account of the fact that each image is affected by various factors such as movement, light, temperature and so on. In this chapter, we apply the proposed time series methods to the SenseCam AUTHOR6DAYS data set (D_1) to illustrate. The results obtained are discussed at the end of the chapter.

3.2 Methods

In what follows, DFA, Equal-time Cross-Correlation Matrix and MODWT are used to explore, respectively, long-range correlations, dynamical changes

and equal-time Cross-Correlation over different time scales for SenseCam non-stationary multivariate time series. The wavelet method also permits examination of details of the eigenvalue spectrum.

3.2.1 Detrended Fluctuation Analysis

DFA, proposed (Peng et al. 1994) is used extensively to estimate the Hurst exponent, (indicating the quantity of long term memory in non-stationary time series data), having the advantage that it can remove different trends due to external factors in the data and reduce noise. We apply a similar approach therefore, to initial analyses of lifelog data streams. The DFA algorithm is illustrated in what follows.

We first generate the images time series $X(t)$. The series $X(t)$ is then divided into boxes of equal size, n. In each box n, a least-squares function, $z(t)$=$at + b$ is fitted to the data representing the trend in that box. The root-mean-square fluctuation of this integrated and detrended time series is calculated by:

$$F(n) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[X(t) - z(t)\right]^2}, \tag{3.1}$$

The calculations are repeated for all considered n. We are interested in the relation between $F(n)$ and size of segment n. In general $F(n)$ will increase with the size of segment n, (Peng et al. 1994).

This linear dependence indicates the presence of self-fluctuations and the slope of the line $F(n)$ determines the scaling exponent H (Peng et al. 1995, Absil et al. 1999, Lee et al. 2002, Stam et al. 2005, Rodriguez et al. 2007). If the

observable $X(t)$ are random uncorrelated variables or short-range correlated variables, the behaviour is expected to obey a power law relationship of the form:

$$F(n) \sim n^H \tag{3.2}$$

The exponent H is called the Hurst exponent and represents the autocorrelation properties of the time series (Acharya et al. 2002, Lee et al. 2002, Stam et al. 2005, Rodriguez et al. 2007, Gifani et al. 2007, Phinyomark et al. 2009) in the following way: if

- $0 < H < 0.5$ then the process has an intermediate memory, and it exhibits anti-correlations (or anti-persistent behaviour), i.e. deviations of one sign are generally followed by deviations with the opposite sign.
- $0.5 < H < 1$ then the process shows evidence of long memory, and it exhibits positive correlations (or persistence), i.e. deviations tend to keep the same sign.
- $H = 0$ then the process has a short memory, corresponding to white noise[14], where fluctuations at all frequencies are equally present.
- $H = 0.5$ then the process is indistinguishable from a random process with no memory.
- $1 < H < 2$ then the process is non-stationary.

[14]In signal processing, white noise is a random signal having equal intensity at different frequencies, giving it a constant power spectral density (Martin 2001)

3.2.2 Correlation Dynamics

The DFA method provides the initial background analysis for lifelogging time series data. The equal-time Cross-Correlation Matrix can be formed and used to characterise dynamical changes in non-stationary multivariate series of this type. Correlation is a typical statistical measure of the strength and direction of a linear relationship between two random variables (Plerou et al. 1999).

The equal-time Cross-Correlation Matrix between time series of images is calculated using a sliding window, where the number of pixels in one image, N, is smaller than the window size T. Given pixels $G_i(t)$, i=1,...,N, of a collection of images, we normalise X_i within each window in order to standardise the different pixels for the images as follows:

$$g_i(t) = \frac{G_i(t) - \overline{G_i(t)}}{\sigma_{(i)}} \tag{3.3}$$

where $\sigma_{(i)}$ is the standard deviation of G_i for image numbers i=1,...,N, and $\overline{G_i}$ is the time average of G_i over a time window of size T. Then the equal-time Cross-Correlation Matrix may be expressed in terms of $g_i(t)$

$$C_{ij} \equiv \left\langle g_i(t) g_j(t) \right\rangle \tag{3.4}$$

The elements of C_{ij} are limited to the domain -1$\leq C_{ij} \leq$1,where C$_{ij}$=$\pm$1 defines perfect positive/negative correlation and C_{ij}=0 corresponds to no correlation. In matrix notation, the correlation matrix can be expressed as $\mathbf{C} = \frac{1}{T}\mathbf{G}\mathbf{G}^t$ where t is the transpose of a matrix and G is an $N \times T$ matrix with elements g_{it}.

The eigenvalues λ_i and eigenvectors $\overline{v}_i$ of the correlation matrix $\mathbf{C}$ are found

from the eigenvalue (characteristic) equation $\mathbf{C}\overline{v_i} = \lambda_i \overline{v_i}$. The eigenvalues are then ordered by size, such that $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N$. Given that the sum of the diagonal elements of a matrix (the *Trace*) remains constant under linear transformation (Schindler et al. 2006), $\sum_i \lambda_i$ must always equal the Trace of the original correlation matrix. Hence, if some eigenvalues increase then others must decrease, to compensate, and vice versa, (a feature known as *Eigenvalue Repulsion*).

There are two limiting cases for the distribution of the eigenvalues: (i) with perfect correlation, $C_i \approx 1$, when the largest is maximised with value N, (all others taking value zero); (ii) when each time series consists of random numbers with average correlation $C_i \approx 0$ and the corresponding eigenvalues are distributed around 1, (where any deviation is due to spurious random correlations). Between these two extremes, the eigenvalues at the lower end of the spectrum can be much smaller than λ_{max}. To study the dynamics of each of the eigenvalues using a sliding window, we normalise each eigenvalue in time using

$$\tilde{\lambda}_i(t) = \frac{(\lambda_i - \overline{\lambda})}{\sigma^{\lambda}} \tag{3.5}$$

where $\overline{\lambda}$ and σ^{λ} are the mean and standard deviation of the eigenvalues over a particular reference period, respectively. This normalisation allows us to visually compare eigenvalues at both ends of the spectrum, even if their magnitudes are significantly different. The reference period used to calculate the mean and standard deviation of the eigenvalue spectrum can be chosen to be a low volatility sub-period (which helps to enhance the visibility of high volatility periods) or can be the full time-period studied.

3.2.3 Wavelet Multiscale Analysis

Wavelet analysis is one of the most important tools for statistical signal extraction, filtering and denoising (Pollock 2006). The advantage of wavelet analysis is the ability to decompose a signal into time-scale (or time-frequency) basis. If we look at a signal in a large 'window', we can see the gross features. Conversely, in a small 'window', detailed features would be apparent. Thus, the wavelet break down of a signal is very useful to show markedly different behaviour in the different time periods.

Specifically, the wavelet transform decomposes a signal into sets of coefficients where each set of coefficients is associated with a spatial scale and each coefficient in a set is associated with a particular location. Each single coefficient is called an 'atom' and the set of coefficients for each scale a 'crystal' (Gallegati 2008). The wavelet coefficients are obtained through the *mother* and *father* wavelets. The former represent the detailed (high frequency) components and the latter represent the smooth (low frequency) components of the signal, respectively.

In particular, the discrete wavelet transform (DWT) is useful in dividing the data series into components of different frequency, so that each component can be studied separately to investigate the series in depth (Burrus et al. 1998). Fig. 3.1 shows a flowchart for a three-level DWT. As seen in Fig. 3.1, the signal is first passed through a low pass filter (L_0). The signal is also decomposed simultaneously using a high-pass filter (Hi). A_3 is the approximation signal which shows the general trend of the low frequency component. D_1, D_2 and D_3 are details of signal and are associated with its high frequency component.

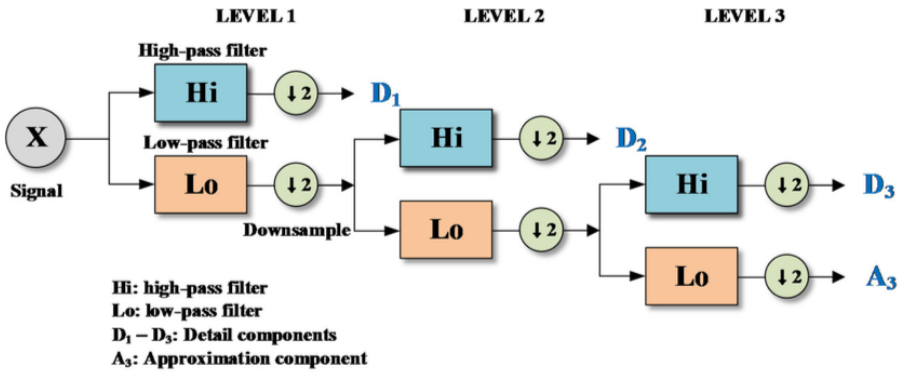

Figure 3.1: Flowchart for a three-level DWT. Source from: *Seo et al. (2017)*

It should be noted, however, that one of the limitations of DWT is a restriction on the length of the data set to a multiple of 2^j. Secondly, the output generated by DWT is highly dependent on the origin of the signal. A small shift in origin affects the outputs generated (Jothimani et al. 2016). Due to these shortcomings, a modification of DWT called Maximum Overlap Discrete Wavelet Transform, (MODWT) is applied.

3.2.3.1 MODWT

The MODWT, is a linear filter that transforms a series into coefficients related to variations over a set of scales (Burrus et al. 1998). Like the DWT it produces a set of time-dependent wavelet and scaling coefficients with basis vectors associated with a location t and a unitless scale τ_j=2^{j-1} for each decomposition level j=1,...,J_0. *Unlike* the DWT, the MODWT has a high level of redundancy. However, it is *non-orthogonal* and can handle any sample size N. MODWT retains downsampled[15] values at each level of the decomposition

[15]Downsampling or decimation of the wavelet coefficients retains half of the number of coefficients that were retained at the previous scale. Downsampling is applied in the Discrete

that would be discarded by the DWT (Percival & Walden 2006). This reduces the tendency for larger errors at lower frequencies when calculating frequency dependent variance and correlations, as more data are available. Fig. 3.2 shows a flowchart for three-level MODWT. As seen in Fig. 3.2, the MODWT decomposes an original signal X into a low-pass filtered approximation component (A_3) and high-pass filtered detail components (D_1, D_2 and D_3).

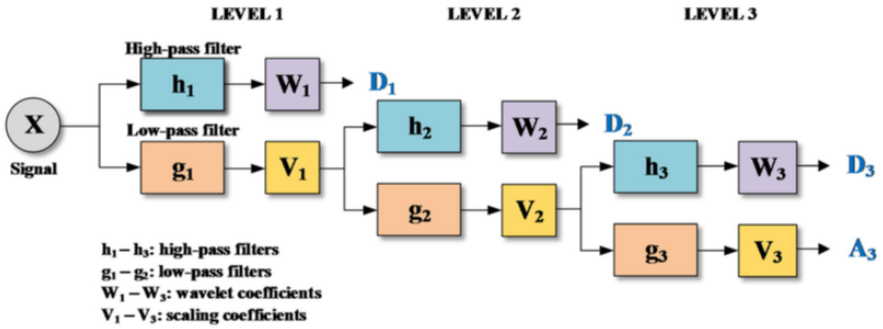

Figure 3.2: Flowchart for a three-level MODWT. Source from: *Seo et al. (2017)*

Decomposing an infinite sequence X_t of Gaussian random variables, using the MODWT to J_0 levels, theoretically involves the application of J_0 pairs of filters. Following Percival & Walden (2006) then, the filtering operation at the j^{th} level consists of applying a rescaled mother (high frequency) wavelet to yield a set of scaling coefficients $\tilde{W}_{j,t}$ and is given by:

$$\tilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l} \tag{3.6}$$

and a rescaled father (low frequency) wavelet to yield a set of scaling coefficients $\tilde{V}_{j,t}$ given by:

Wavelet Transform.

$$\tilde{V}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l} \tag{3.7}$$

for all times $t = ..., -1, 0, 1, ...$; where the MODWT wavelet and scaling filters $\tilde{h}_{j,l}$ and $\tilde{g}_{j,l}$ are obtained by rescaling the DWT filters as follows:

$$\tilde{h}_{j,l} = \frac{h_{j,l}}{2^j} \tag{3.8}$$

and

$$\tilde{g}_{j,l} = \frac{g_{j,l}}{2^j} \tag{3.9}$$

where wavelets for the j^{th} level are a set of scale-dependent localised differencing and averaging operators and can be regarded as rescaled versions of the originals. The j^{th} level equivalent filter coefficients have a width $L_j = (2^j - 1)(L-1)+1$, where L is the width of the $j = 1$ base filter. In practice the filters for $j > 1$ are not explicitly constructed because the detail and scaling coefficients can be calculated, using an algorithm that involves the $j = 1$ filters operating recurrently on the j^{th} level scaling coefficients, to generate the $j+1$ level scaling and detail coefficients.

The MODWT offers several advantages over the DWT:

- The MODWT can handle all sample sizes N; For complete decomposition of J levels is achievable the DWT requires N to be a multiple of 2^J;
- The MODWT is not influenced by circular shifting of the signal (the signal does not change the pattern of wavelet transform coefficients), whereas DWT depends upon the starting point of the signal;
- The MODWT wavelet coefficients increases the effective degrees of free-

dom on each scale and thus produces a more asymptotically efficient wavelet variance estimator than DWT modestly (Cornish et al. 2006, Gallegati 2008);

3.2.3.2 Wavelet Variance

The decomposition of a time series into scale-dependent coefficients using the MODWT permits the statistical analysis of a signal as a function of scale. The MODWT is energy-conserving; i.e.

$$\| X \|^2 = \sum_{j=1}^{J_0} \| \tilde{W}_j \|^2 + \| \tilde{V}_{J_0} \|^2 \tag{3.10}$$

Where X is as above and $\tilde{W}$ and $\tilde{V}$ are defined in Equations (3.6), (3.7), respectively. The wavelet variance $\nu_X^2(\tau_j)$ is defined as the expected value of $\tilde{W}_{j,t}^2$ if we consider only the non-boundary coefficients[16]. An unbiased estimator of the wavelet variance is formed by removing all coefficients that are affected by boundary conditions and is given by

$$\nu_X^2(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^2 \tag{3.11}$$

where $M_j = N - L_j + 1$ is the number of non-boundary coefficients at the j^{th} level (Percival & Walden 2006). The wavelet variance decomposes the variance of a process on a scale-by-scale basis (at increasingly higher resolutions of the signal) and allows us to explore how a signal behaves over different time horizons.

[16]The MODWT treats the time-series as if it were periodic using "circular boundary conditions". There are L_j wavelet and scaling coefficients that are influenced by the extension, and which are referred to as the boundary coefficients.

3.2.3.3 Wavelet Covariance and correlation

The wavelet covariance between two functions $X(t)$ and $Y(t)$ is similarly defined to be the covariance of the wavelet coefficients at a given scale. The unbiased estimator of the wavelet covariance at the j^{th} scale is given by:

$$\nu_{XY}(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^{X(t)} \tilde{W}_{j,t}^{Y(t)} \tag{3.12}$$

where all the wavelet coefficients affected by the boundary are removed, and $M_j = N - L_j + 1$.

The MODWT estimate of the wavelet cross-correlation between functions $X(t)$ and $Y(t)$ may be calculated using the wavelet covariance and the square root of the wavelet variance of the functions at each scale j. The MODWT estimator of the wavelet correlation is given by:

$$\rho_{XY}(\tau_j) = \frac{\nu_{XY}(\tau_j)}{\nu_X(\tau_j)\nu_Y(\tau_j)} \tag{3.13}$$

where, at scale j, $\nu_{XY}(\tau_j)$ is the covariance between $X(t)$ and $Y(t)$, $\nu_X(\tau_j)$ is the variance of $X(t)$ and $\nu_Y(\tau_j)$ is the variance of $Y(t)$.

3.3 Results

In this chapter, we applied methods discussed in Section 3.2 to the AUTHOR6DAYS data set (D_1) (Section 2.6). A total of 10,260 SenseCam images[17] were available in the time series.

[17]Generated by author over 6 days including a weekend (Saturday to Thursday)

3.3.1 Detrended Fluctuation Analysis

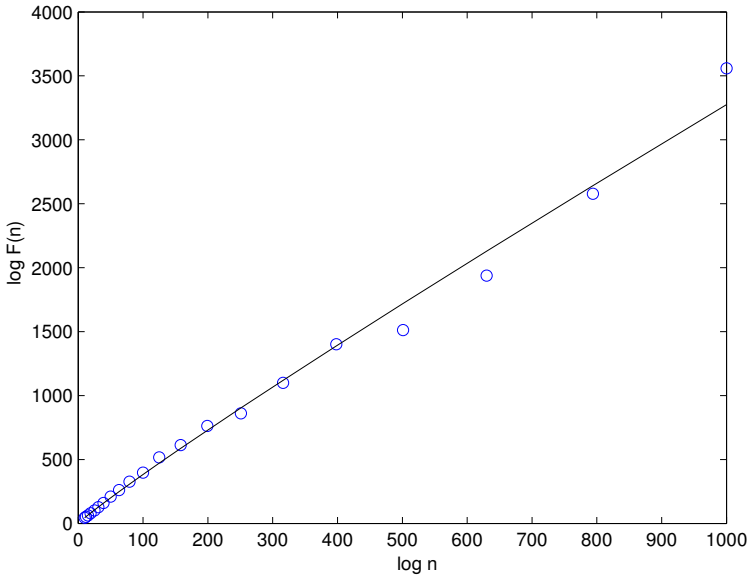

Figure 3.3: Plot of log $F(n)$ vs log n , where n = box size, ranging from 10 to 1000 for the AUTHOR6DAYS data set (D_1).

According to Section 3.2.1, if the scaling exponent H= 0.5, there is no correlation and the time series is uncorrelated; if H> 0.5, the time series exhibits positive correlation and if H> 0.5, the time series is anti-correlated. In Fig. 3.3, long-range correlation is demonstrated for the AUTHOR6DAYS data set D_1 with an exponent H=0.93 Equation (3.2). This indicates that the time series is not a random walk[18], but is cyclical, implying that continuous low levels of background information are picked up constantly by the device. Consequently, the DFA provides a measure of the predominance of 'typical' backgrounds or environments.

[18]A random walk is a mathematical formalisation of a path that consists of a succession of random steps.

3.3.2 Dynamics of the largest Eigenvalue for different sliding window sizes

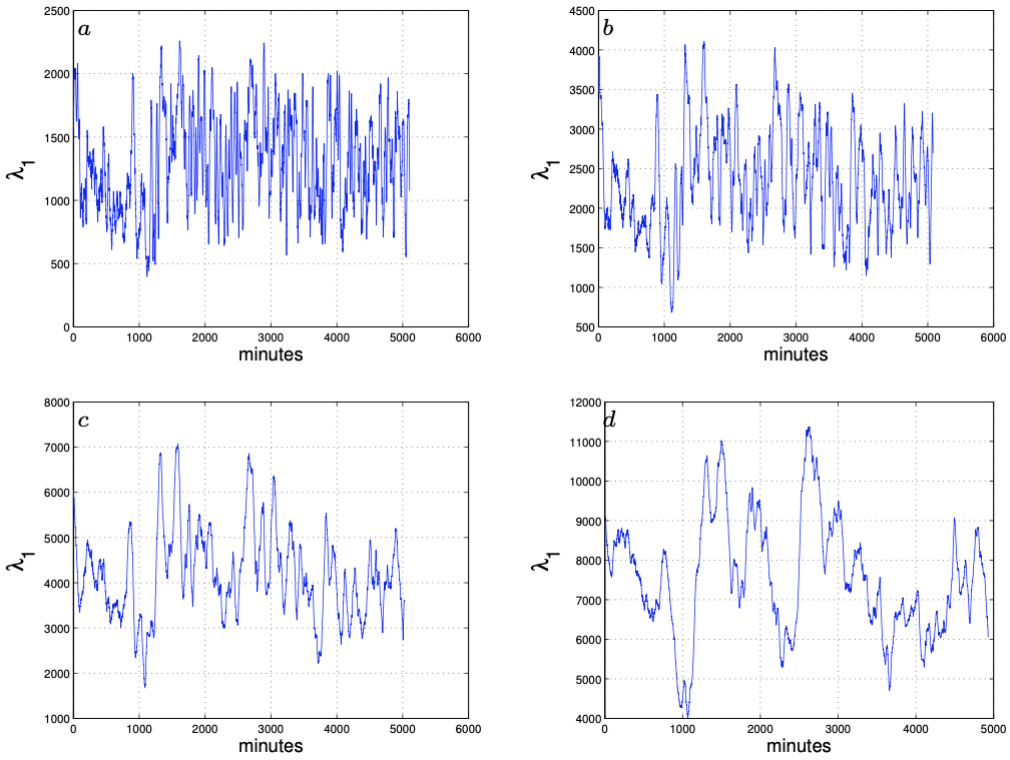

Figure 3.4: The largest Eigenvalue Distribution using a sliding window of 50 Images(a), 100 Images(b), 200 Images(c) and 400 Images(d) of the AUTHOR6DAYS data set (D_1).

Fig. 3.4 shows the time series of the Largest Eigenvalue dynamics across different wavelet scales. From these, we note the following features:

- At increased scale, (i.e for longer time periods), of the wavelet crystal components, increased smoothing was observed- as expected. Increasing the scale has the effect of removing some of the high frequency small-scale changes, typically associated with noise.

- The different features, found at various scales, suggest that the correlation matrix captured different major events with different time horizons. This will be examined in more detail in the next subsection.

3.3.3 Wavelet Multiscale Anlaysis

In order to perform wavelet analysis on a time series, we first need to consider the selection of a wavelet filter. There are several available, including the haar (discrete), symmlets and coiflets (symmetric), daublets (asymmetric), etc (as shown in Fig. 3.5), where these differ in terms of characteristics of the in-transfer function and filter length. Daubechies (1992) developed a family of compactly supported wavelet filters of various lengths, where the least asymmetric family or (LA) are defined in even widths with the optimal filter width dependent on the characteristics of the signal and the length of the data series. The filter width chosen for this study was the LA8, (where 8 refers to the width of the scaling function). This choice enables accurate calculation of wavelet correlations to the 10^{th} scale, which is appropriate given the length of data series available[19]. Although MODWT can accommodate any level, J_0, the largest level is chosen in practice so as to prevent decomposition at scales longer than the total length of the data series (hence the choice of the 10^{th}) while still containing enough detail to capture subtle changes in the signal (Percival & Walden 2006). The MODWT was implemented using the WMTSA (Wavelet Methods for Time Series Analysis)[20] Toolkit for Matlab (Cornish et al. 2003).

[19]This implies 10^{th} is largest possible since the j^{th} level equivalent filter coefficients have a width $L_j = (2^j - 1)(L - 1) + 1$.

[20]https://atmos.washington.edu/wmtsa/

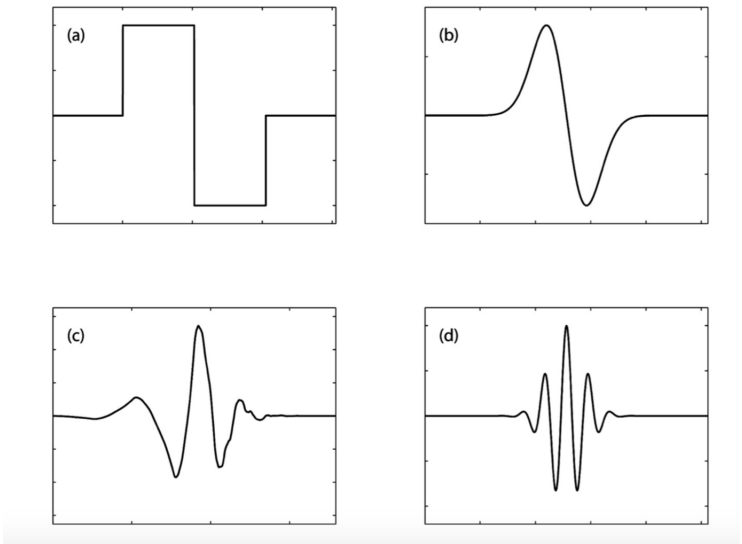

Figure 3.5: Examples of wavelet filters (a) Haar (b) Daublets (c) Coiflets (d) Symmlets. Source:*(Zubal' 2015)*

First, the MODWT of the pixels for each image was calculated within each window and the correlation matrix between pixels at each scale found. The eigenvalues of the correlation matrix in each window were determined, and the eigenvalue time series was normalised in time. Then the largest eigenvalue for different window sizes was analysed for the AUTHOR6DAYS data set (D_1). Results are illustrated by a heatmap (Fig. 3.6) and discussed below.

3.3.3.1 Dynamics of the largest Eigenvalue at various wavelet scales

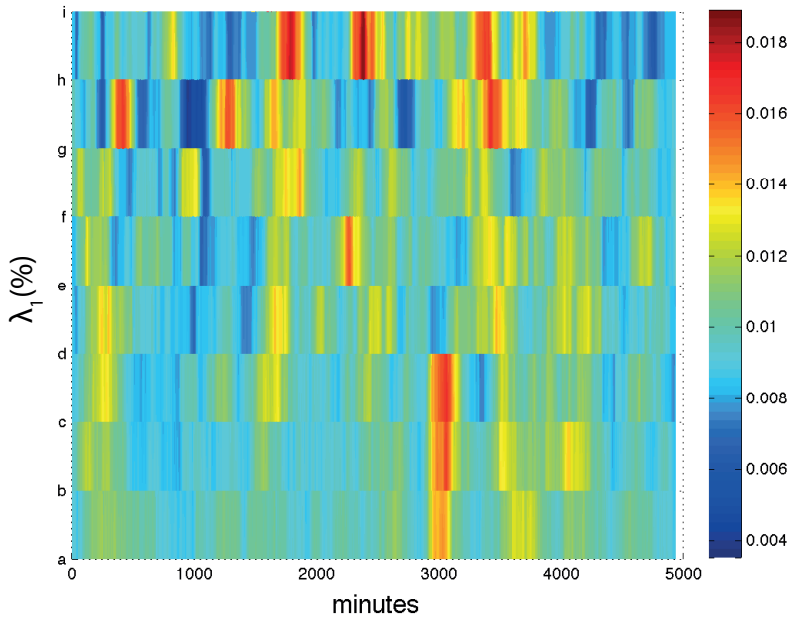

Figure 3.6: Heatmap diagram of the largest Eigenvalue λ_1 dynamics across 9 wavelet scales of the AUTHOR6DAYS data set (D_1). Scales 1 *(a)* to 9 *(i)* correspond to time periods of 1-2, 2-4, 4-8, 8-16, 16-32, 32-64, 64-128, 128-256 and 256-512 minutes, respectively

Fig. 3.6 shows time series dynamics of the largest Eigenvalue across different wavelet scales. Some peaks are consistently captured by the SenseCam at certain scales, such as the peak around 3000 minutes, (captured by wavelet scales 1, 2, 3 and corresponding to a 1-2 minute period, a 2-4 minute period and a 4-8 minute period, respectively). These peaks should help to identify *major events* or *activities* in the data. The different features, found at various scales, suggest that the correlation matrix captured different major events with different time horizons.

3.3.3.2 The largest Eigenvalue λ_1 compared with the ratio of λ_1/λ_2 dynamics

We also wished to determine whether the sub-dominant eigenvalues e.g. λ_2 hold further information on the key sources of major events and what information these contribute additionally from the images. In Fig. 3.7, the dynamics of the series from the MODWT analysis for the largest eigenvalue and changes of the eigenvalue ratio λ_1/λ_2 were examined. Here, we detail several scenarios for the peaks in the largest eigenvalue and the ratio of the largest to the next largest eigenvalue for a window size of 400 images.

Analysis of Scenarios: We have studied both the largest eigenvalue λ_1 and the ratio of λ_1/λ_2 time series for a window size of 400 images to try to identify the position and nature of such events from the real images generated from SenseCam. The different features, found at various scales, suggest that the correlation matrix captured different major events with some features consistent and others specific to certain scales. We group peaks in reporting, where scenarios are very similar, with more details as given below.

1. Peaks a_1, b_1, c_3

This group of fluctuations in the signal relate to the subject arriving at the office, and switching on the lights and the laptop. The laptop colour is white, with the screen the largest object in the field of view of the SenseCam. Thus, the lights and the laptop mainly contribute to these peaks.

Peak a_1 was less noticeable for λ_1/λ_2 due to an increase in λ_2. The subject was moving during this event. Peak a_1 corresponds to the wavelet scale of a 1-2 minute period. Over this short time period, λ_2 carries other information on the environment, such as the white wall, desk and chairs in the office (peak

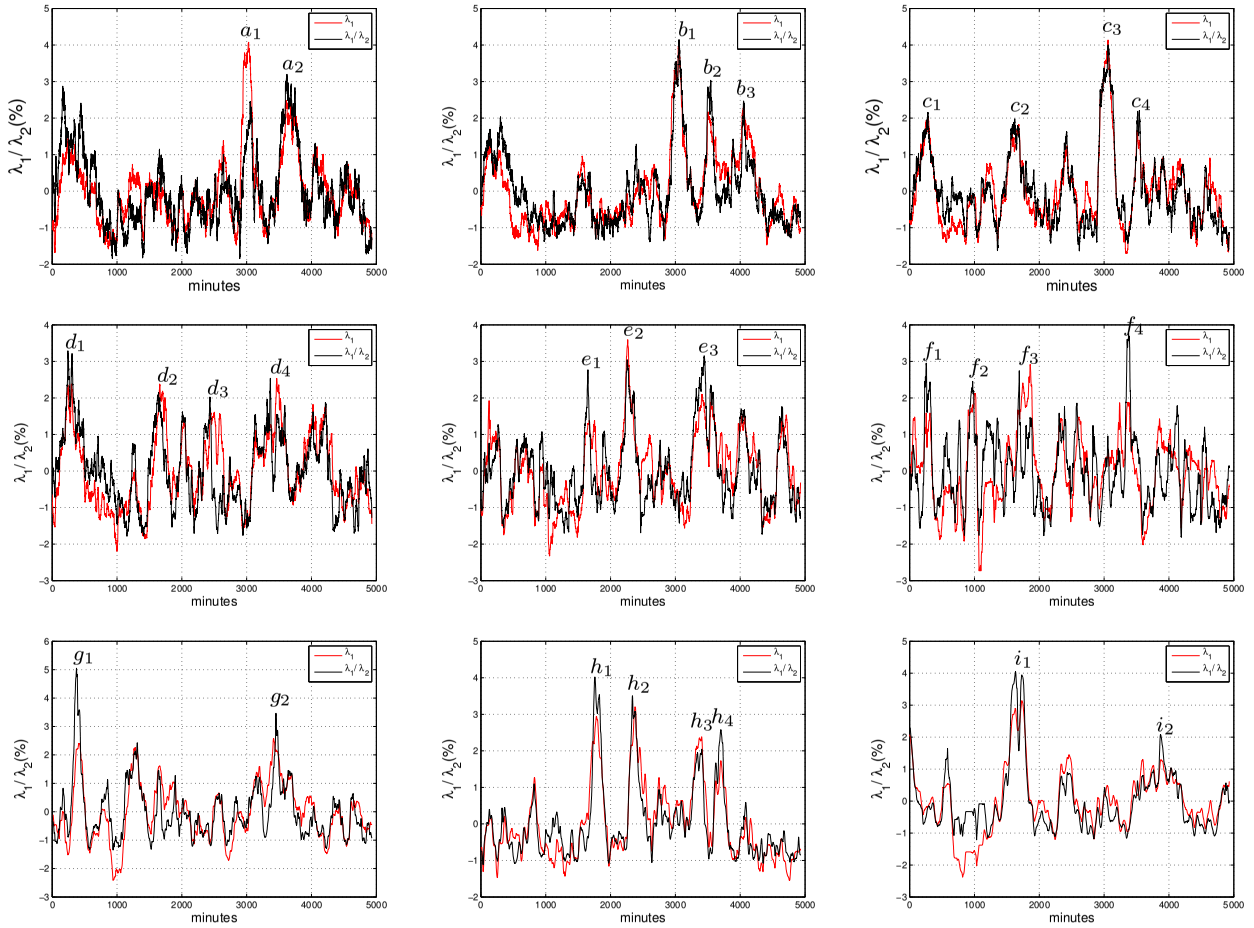

Figure 3.7: The Largest Eigenvalue λ_1 *(red)* and ratio λ_1/λ_2 *(black)* dynamics across 9 wavelet scales of the AUTHOR6DAYS data set (D_1). *(a-c)* are for scales 1 to 3 (1-2, 2-4 and 4-8 minutes periods, respectively), *(d-f)* are for scales 4 to 6 (8-16, 16-32 and 32-64 minute periods, respectively) and *(g-i)* are for scales 7 to 9 (64-128, 128-256 and 256-512 minute periods, respectively.)

a_1). As the time period increases, the subject stopped moving and sat down so that the λ_1 and λ_1/λ_2 peaks coincide at peak c_3.

2. Peaks a_2, b_2, b_3, c_4

These peaks occurred when the subject was working in front of the laptop, with slight movements, partially obscuring the camera, (e.g. hair or hand interruption). Here, peaks are due to the differences in light level from the laptop, ceiling lights and another desktop. Thus, the laptop screen and ceiling light register higher pixel values than other objects, for which changes are picked up by the SenseCam.

3. Peaks c_1, d_1, f_1, g_1

These peaks refer to the period while the subject was walking in the city centre during the day time. During this period, the subject visited several shops, with lights in the shops reflecting higher pixel values than other objects that contribute to the peaks. A point to note was that a small dip in peaks d_1 and f_1 can be observed, corresponding to the subject picking up clothes, i.e. introducing a new major object into the image field at that time.

In peaks d_1, f_1, g_1, the ratio λ_1/λ_2 implies that λ_1 increased with λ_2. The scenario involved the subject moving along a street, visiting shops, with the strong light in the shops dominating during less movement and less change in other objects.

4. Peaks c_2, d_2, e_1, f_3, h_1, i_1

These peaks involved the subject walking from home to office in the morning, working in front of the laptop and talking with her colleagues. Note that the office was dark until the subject switched on the lights and laptop, consequently the lights and laptop introduction are highlighted by the peaks. While

the subject was standing and talking with colleagues, the camera captured office lights which contributed to the peaks.

At peak f_3 in wavelet scale 6, corresponding to a 32-64 minute period, the highest value of the λ_1/λ_2 ratio occurs earlier in time than that for the largest eigenvalue λ_1, which may imply capture of other effects compared to that of the largest eigenvalue λ_1; This shift in the peak of λ_1 relative to that in λ_1/λ_2 occurred when the subject was talking with colleagues (so during localised position change relative to lighting). These movements may contribute to difference in peak position.

For peaks h_1, i_1, the ratio λ_1/λ_2 was less affected, implying that λ_1 increased with λ_2. These features involved the subject sitting in front her desk and standing in front of a colleague's desk, talking to colleagues. During the standing period, more objects were captured by the SenseCam especially the lights, with light changes also featuring during localised movement, which may have caused increases in λ_1 and λ_2.

5. Peak d_3

The d_3 peak ratio of λ_1/λ_2 occurs slightly earlier in time than that for the largest eigenvalue λ_1, corresponding to a period when the subject was sitting in front of her laptop, with the exception of a short period when the subject moved to another PC on the desk. This produced a small dip in the largest eigenvalue λ_1.

6. Peaks d_4, e_3

These involved the subject sitting in front of her laptop, leaving to have lunch in the cafeteria with colleagues and then returning to the office. The static period in front of the laptop is reflected in a higher pixel value for this

peak. When the subject was moving the images were less clear, hence a small dip around this peak was observed.

For peak d_4, the maximum value for λ_1/λ_2 ratio occurs slightly earlier in time than that for the largest eigenvalue λ_1, which may be caused by the subject changing position from sitting to moving.

7. Peaks e_2, h_2

These describe the movement of the subject from outside coming into the office and sitting in front of the laptop. The white laptop and screen are the major objects in these images.

8. Peak f_2

The peak of the ratio of the λ_1/λ_2 series is shifted slightly in time in comparison to that for the largest eigenvalue λ_1. The peak involved the subject walking down an urban street during the evening. The camera captured the lights on the road and in the shops.

9. Peaks f_4, g_2, h_3

These peaks occurred while the subject was sitting in front of the laptop; it is important to note that the camera was inadvertently blocked by the subject on numerous occasions.

10. Peak h_4

The subject was sitting in front of the laptop and then visited the cafeteria with colleagues. While there, she ordered lunch, after which she returned to the office and sat in front of the laptop. Some partial or total blocking of the laptop occurs during the seated periods.

11. Peak i_2

This peak was a typical case for the subject's activities. She was working

in front of laptop, with the camera sometimes obscured and sometimes fully capturing the ceiling light.

MODWT gives a clear picture of the movements in the image time series by reconstructing them using each wavelet component. The method captured the features markedly apparent at specified scales. A number of features from the image are reproduced and can be examined by studying these eigenvalue series.

For lower scales (1-8 minutes), most of the peaks highlight dramatic light level changes, such as the subject switching on the lights in the dark office or the camera being totally blocked, etc. The subject activity changes, where more people were involved in an event such as (i) visiting the cafeteria and subsequently returning to the office (about 1 hour) or (ii) travelling from home to the office (about 15 minutes), were highlighted by the middle wavelet scales (8 minutes -1 hour). We found that most distinct events or activities were marked in wavelet scale4 (8 minutes -16 minutes), e.g. in peak d_2, which describes the journey for the subject from the home to the office, (about 15 minutes), and is sufficiently fine-grained to pick up other activities. At higher scales (larger periods 1- 8.5 hours), the SenseCam mostly captured the subject maintaining a single activity, which may generally be of less interest to wearers or analysts due to these being remembered more easily.

For the majority of images, the subject was seated in front of her laptop, with laptop, lights and seating position unchanged over on extended period, contributing consistently high pixel values in a sequence of images. This typical case was always marked by a peak in the SenseCam signal. The signal fluctuation is caused by light level changes, such as the subject moving from

indoors to outdoors, the subject changing position from sitting to moving, movement increase and more people joining in the scene. Note that a movement or multiple person interactions can be captured by specific scales, using the MODWT method. The ratio analysis strongly reinforces observations on the largest eigenvalue over time. The ratio of λ_1/λ_2 has smaller variation compared to that for the largest eigenvalue λ_1. This implies that the second largest eigenvalue (λ_2) carries additional information to describe events, but does not decide the occurrence of major events for SenseCam. It appears, however, that is does carry information for events surrounding the major ones, e.g. possible lead-in, lead-out. Nevertheless it does represent a significant finding.

3.3.4 Evaluation

Video segmentation is useful for video classification, summarization, indexing and retrieval. Extensive research has been done by the computer vision research community to create different approaches and algorithms for video segmentation. Nevertheless, it is still difficult to judge which approach or algorithm works best. Two main reasons can be identified for that: First, there is a lack of a benchmark data sets and a standardized evaluation methodology. There are some common approaches for evaluating video segmentation that focus on Precision/Recall, F1 measure, Receiver Operator Characteristics (ROC), Jacard Coefficient. The boundary precision/recall is the most popular evaluation for image segmentation. However, this metric is of limited use in a video segmentation benchmark, as it evaluates every frame independently. The volume precision-recall metric has been proposed for evaluating video segmentation (Galasso et al. 2013). In this chapter, we focus on the analysis of an

image data set generated by SenseCam, so precision, P, and recall, R metrics are employed for evaluating the performance at different wavelet scales.

We evaluate the performance at the different wavelet scales using the precision[21], P, and recall[22], R, metrics - as defined below. Moreover, we compute the F_1 score[23] as a measure of a method's accuracy (Perry et al. 1955).

$$Precision = \frac{|\text{determined boundaries}| - |\text{wrong boundaries}|}{|\text{determined boundaries}|} \tag{3.14}$$

$$Recall = \frac{|\text{detected reference boundaries}|}{|\text{determined boundaries}|} \tag{3.15}$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \tag{3.16}$$

Table 2.1 in Chapter 2 shows more than 60 ground truth events manually segmented by a user. In order to determine accurate boundaries, each peak point boundary is calculated, (for the difference between neighbouring left- and right- most trough values) (Doherty & Smeaton 2008). This is obviously a crude boundary designation; All values within a peak area are combined so that a signal value is less informative. Significant peaks are determined (distinct events or activities) by λ_1/λ_2 percentage pixel values that are larger than zero.

[21]Precision is the probability that a (randomly selected) retrieved document is relevant.

[22]Recall is the probability that a (randomly selected) relevant document is retrieved in a search.

[23]A measure that combines precision and recall is the harmonic mean of precision and recall.

Table 3.1: Precision, Recall and F_1 measures for MODWT method

	λ_1/λ_2		
Wavelet Scales (minute period)	*Precision*	*Recall*	F_1
1	0.3929	0.4058	0.3992
2	0.7857	0.2029	0.3225
3	0.5000	0.3188	0.3894
4	0.4783	0.3333	0.3929
5	0.5238	0.3043	0.3850
6	0.5789	0.2754	0.3732
7	0.7333	0.2174	0.3354
8	0.9167	0.1739	0.2924
9	1	0.1594	0.2750

Table 3.1 shows the precision, recall and F_1 measure for λ_1/λ_2 at different wavelet scales. As we can see, the wavelet scale 1 gives the best result for the F_1 measure. However, in real life, we argue that middle wavelets 4 to 6 (corresponds to 8-16, 16-32 and 32-64) captures the information that should be of interest to the users. Imagine, for example, the scenario where elderly people have a problem to remember if they have already taken their pills. Such an event normally lasts about 30-60 minutes as people usually take their pills at a certain time of the day. In Table 3.1, overall, most scales return high precision and very low recall. The main weakness as well as strength for wavelet scales is that *different* scales highlight *different distinct events* dependent on the time horizons. Some events at certain scales will be missed, so that the *overall recall*

values are low for this approach. In addition, some activities, such as working in front of the laptop, last for several hours. In manually segmenting 69 events out of 10,260 images only, the detection probability for a given event is quite low. In consequence this approach is still quite crude and we would suggest that further modifications are needed, such as incorporating other than peak distance and weighting scale combinations.

3.4 Summary

In this chapter we performed DFA on image time series recorded by the SenseCam. The results suggest strong long-range correlations in the time series, which means that some information is always picked up by the device, even during relatively static periods. Consequently, DFA provides a useful background summary.

Using the Maximum Overlap Discrete Wavelet Transform (MODWT) estimates of wavelet variance we provide a scale-based analysis of variance that allow us to identify the main event captured by the SenseCam as well as additional information surrounding the main event. The main scale-by-scale results deriving from the MODWT analysis variance are: i) we identified light level as a major event delineator during static periods of image sequences; ii) we have shown that the wavelet scale from 8 minutes to 16 minutes can be seen as the most important time horizon for the identification of distinct events or activities of the users, e.g., the subject changing position from sitting indoors to walking outdoors, as graining is sufficiently fine to enable recording of short term activities. This time scale, on the other hand, is not so fine-grained as to be swamped by noise. The MODWT method provides a powerful tool for

the examination of the nature and quality of the captured SenseCam data for categories of users, such as the case described.

Chapter 4

Random Matrix Theory

4.1 Introduction

In the last Chapter, we have shown that SenseCam image time series exhibit a strong long-range correlation, concluding that the time series does not have the form of a random walk, but is cyclical, with continuous low levels of background information picked up constantly by the device. Further, we adopted a Cross-Correlation Matrix method to highlight *key episodes*, thus identifying natural boundaries between different daily events. However, the length of time series available to estimate the empirical Cross-Correlation Matrix is limited, so that, we speculate, the Cross-Correlation Matrix contains much which corresponds to "random" contributions (Wigner 1951, Dyson 1962, Dyson & Mehta 1963, Mehta & Dyson 1963). As a consequence, this technique results in the identification of a high percentage of noise or routine events, with its application in many disciplines of science (Sengupta & Mitra 1999, Mehta 2004, Mezzadri et al. 2005, Couillet & Debbah 2011).

Random Matrix Theory (RMT) was first introduced by author Wigner

(1951) to explain the energy levels of complex interactions between nuclei interactions. This technique has subsequently been applied to noise filtering in financial time series (Laloux et al. 2000, Plerou et al. 2002). In this chapter, we investigate whether RMT can be used to distinguish routine events from important events in the SenseCam series. Our goal is to segment the content of the Cross-Correlation Matrix into two: (a) the part that conforms to the properties of Random Matrix ('noise') and (b) the part that deviates from random (i.e. has 'information' on important events).

4.2 Methods

In this section, we analyse the distribution of the correlation coefficients of the Cross-Correlation Matrix **C** to examine if it exhibits non-Gaussian fluctuations with fat tails and long-range time correlations. Next, we apply RMT methods to analyse the **C** to see if the eigenvalues agree with RMT predictions. Further, we use the Inverse Participation Ratio (IPR)[24] concept to analyze the eigenvectors of the Cross-Correlation Matrix to find out if the eigenvalue spectrum of **C** deviates from RMT prediction.

4.2.1 Random Matrix Theory

As described by Laloux et al. (2000), Plerou et al. (2002) and others, the Random Matrix **R** can defined by:

$$R = \frac{1}{T} A A^t \tag{4.1}$$

[24]Inverse participation ratio is a measure of localization and is defined as the average of the absolute value of the fourth power of the wave function.

Where $\mathbf{A}$ is an $N \times T$ matrix whose elements are i.i.d.[25] random distributed, with zero mean and unit variance. In particular, the limiting property for the sample size $N \to \infty$ and sample length $T \to \infty$, (providing that the 'quality factor' $Q = T/N \geq 1$ is fixed (Bouchaud & Potters 2003)), has been analysed for the distribution of eigenvalues λ of the Random Matrix $\mathbf{R}$ (Sengupta & Mitra 1999), given by:

$$P_{rm}(\lambda) = \begin{cases} \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} & \text{for } \lambda_- \leq \lambda_i \leq \lambda_+ \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

where λ_- and λ_+ are the minimum and maximum eigenvalues of R, respectively, given by

$$\begin{cases} \lambda_+ = \sigma^2(1 + \frac{1}{Q} + 2\sqrt{\frac{1}{Q}}) \\ \lambda_- = \sigma^2(1 + \frac{1}{Q} - 2\sqrt{\frac{1}{Q}}) \end{cases} \tag{4.3}$$

Then σ^2 is the variance of the elements of $\mathbf{G}$ (defined in Chapter 3 section 3.2.2 above) and λ_- and λ_+ are the bounds of the theoretical eigenvalue distribution. Eigenvalues that fall outside this region are said to deviate from the expected values of the Random Matrix. Hence, by comparing the empirical distribution of the eigenvalues of the Cross-Correlation Matrix to the distribution for a Random Matrix, as given in Equation (4.2), we can identify those key eigenvalues which can be used to identify the specific information relating to the system. Eigenvector analysis enables identification of the specific information present, in terms of contributory components.

[25] i.i.d $\equiv$ independent and identically distributed

4.2.2 Eigenvector Analysis

Differences between the eigenvalues $P(\lambda)$ of $\mathbf{C}$ and RMT eigenvalues, $P_{rm}(\lambda)$ should also be evident, therefore, in the statistics of the corresponding eigenvector components. In order to interpret this deviation of the eigenvectors, we note that the largest eigenvalue is an order of magnitude larger than the others, which constrains the remaining $N-1$ eigenvalues, since the trace of $\mathbf{C}$, $Tr[C]$ sums to N. Hence, in order to analyse the contents of the remaining eigenvectors, we must first remove the effect of the largest eigenvalue. To do this we can use linear regression (Plerou et al. 2002):

$$G_i(t) = \alpha_i + \beta_i G^{large}(t) + \epsilon_i(t) \tag{4.4}$$

where

$$G^{large} = \sum_{1}^{N} u_i^{large} G_i(t) \tag{4.5}$$

and N is the number of images in our sample. Here u_i^{large} corresponds to the components of the largest eigenvector. The Cross-Correlation Matrix $\mathbf{C}$, is then recalculated, using the residuals $\epsilon_i(t)$. If we quantify the 'remainder variance', (i.e., that part not 'explained' by the largest eigenvalue) as $\sigma^2 = 1 - \lambda_{large}/n$, this value can be used to recalculate our values of $\lambda_{\pm}$.

To assess whether random effects are less marked further away from the RMT upper boundary λ_+, we use the Inverse Participation Ratio (IPR). The IPR allows quantification of the number of components that participate significantly in each eigenvector and tells us more about the level and nature of the deviation from RMT. The IPR of the eigenvector u^k is given by $I^k \equiv \sum_{l=1}^{N} [u_l^k]^4$ and allows us to compute the inverse of the number of eigenvector components

that contribute significantly to each eigenvector.

4.3 Results

In this chapter, we analysed the AUTHORTYPICAL data set (D_2) introduced in Section 2.6[26]. The author recorded an average day of her life: commuting to the office in the morning, sitting and working in the office at a desk, talking with colleagues and sharing lunch in the cafeteria, as well as commuting back home in the evening and so on. When exploring one's lifelog, reviewing routine or 'boring' events has only limited interest, depending on the device's purpose. Efforts to determine automatically which events are most important or unusual (e.g., talking with a colleague as opposed to working in front of a computer), is an open research challenge. In order to offer a preliminary attempt to distinguish routine or 'boring' events from important events, we apply the RMT method which has proved useful in noise filtering for the Cross-Correlation Matrix enabling important signal features to be highlighted.

4.3.1 Statistics of Correlation Coefficients

In order to quantify correlations, we first analyse the distribution $P(C_{ij})$ of the elements $\{C_{ij} : i = j\}$ of the Cross-Correlation Matrix $\mathbf{C}$ and distribution $P(R_{ij})$ of the elements $\{R_{ij} : i = j\}$ of the Random Matrix. Fig. 4.1 shows that $P(R_{ij})$ appears to be consistent with a Gaussian with zero mean, in contrast to $P(C_{ij})$, which is asymmetric, with a longer positive tail and a high peak, implying that positively correlated behaviour is more prevalent than

[26]The author wearing SenseCam generated total 2096 images for her a typical working day

negative. We argue that the fat positive tail represents significant or unusual events in the data stream. In addition, we see that much of $P(C_{ij})$ falls within the Gaussian curve for the control, suggesting the possibility that observed similarities with **R** in the Cross-Correlation Matrix **C** may be an effect of randomness.

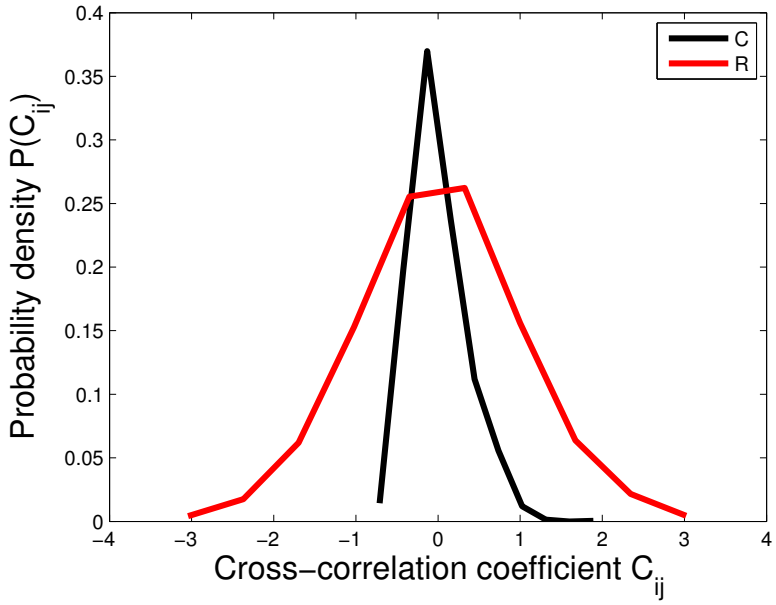

Figure 4.1: Correlation Coefficient Distribution of the AUTHORTYPICAL data set (D_2) for the Cross-Correlation Matrix **C** for SenseCam data (black) and Random Matrix **R** (red).

4.3.2 Eigenvalue Analysis

Given the aim is to distinguish between information (major events) and noise in the Cross-Correlation Matrix **C**, we compare the eigenvalue distribution $P(\lambda)$ of **C** with the corresponding eigenvalue distribution predicted by RMT $P_{rm}(\lambda)$. We compute the eigenvalues λ_i of **C**, where λ_i are rank ordered ($\lambda_{i+1}>\lambda_i$). Fig. 4.2 compares the probability distribution $P(\lambda)$ with $P_{rm}(\lambda)$.

We note the presence of a well-defined 'bulk' of eigenvalues which fall within the bounds $[\lambda_-,\ \lambda_+]$ for $P_{rm}(\lambda)$. We also note deviations for ($\approx 80\%$) largest and smallest eigenvalues. This suggests that the Cross-Correlation Matrix has captured most major events from the data streams, but still contains some percentage of noise ($\approx 20\%$).

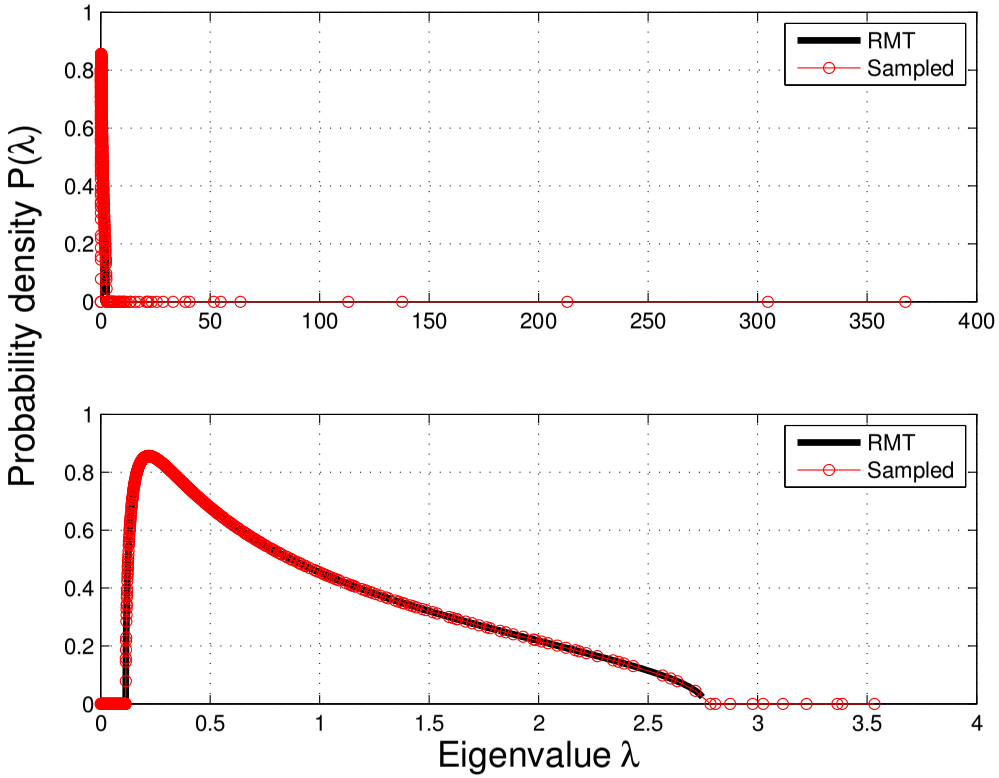

Figure 4.2: Eigenvalue Distribution of the AUTHORTYPICAL data set (D_2) for the Correlation Matrix $\mathbf{C}$ for SenseCam data: (a) Full and (b) Partial spectral distribution.

4.3.3 Eigenvector Analysis

Deviations from RMT: The deviations of $P(\lambda)$ from the RMT result $P_{rm}(\lambda)$ suggest these deviations will also be observed in the statistics of the corresponding eigenvector components (Conlon et al. 2009). The distribution of the components $\{u_l^k; l = 1, ..., N\}$ of eigenvector U^k of a Random Matrix $\mathbf{R}$ should conform to a Gaussian distribution with zero mean and unit variance. First, we consider the distribution of eigenvector components of $\mathbf{C}$. We analyse $P(u)$ for $\mathbf{C}$ and choose one typical eigenvalue λ_k from the bulk ($\lambda_- \leq \lambda_k \leq \lambda_+$) defined by $P_{rm}(\lambda)$ from Equation (4.2). This aims to demonstrate that $P(u)$ for a typical U^k from the bulk shows reasonable total agreement with the RMT result $P_{rm}(u)$. Other eigenvectors, belonging to eigenvalues within the bulk, yield consistent results, (in agreement with those of the previous sections) to Random Matrix predictions. We test the agreement of the distribution $P(U)$ with $P_{rm}(u)$ by calculating the kurtosis, which for a Gaussian has the value 3. In Fig. 4.3, we find that the largest eigenvector (≈ 4.07) deviates significantly from the Gaussian value. The second and third Eigenvectors (≈ 3.7, 3.8) are also have $P(U)$ larger than the Gaussian value, but not markedly so. The Eigenvector from the bulk is however consistent with the Gaussian value 3. These findings suggest that the *largest eigenvalue* (corresponding to the *largest eigenvector*) represents information from the image on the largest change in the SenseCam recording, which is typically the illumination level or light level.

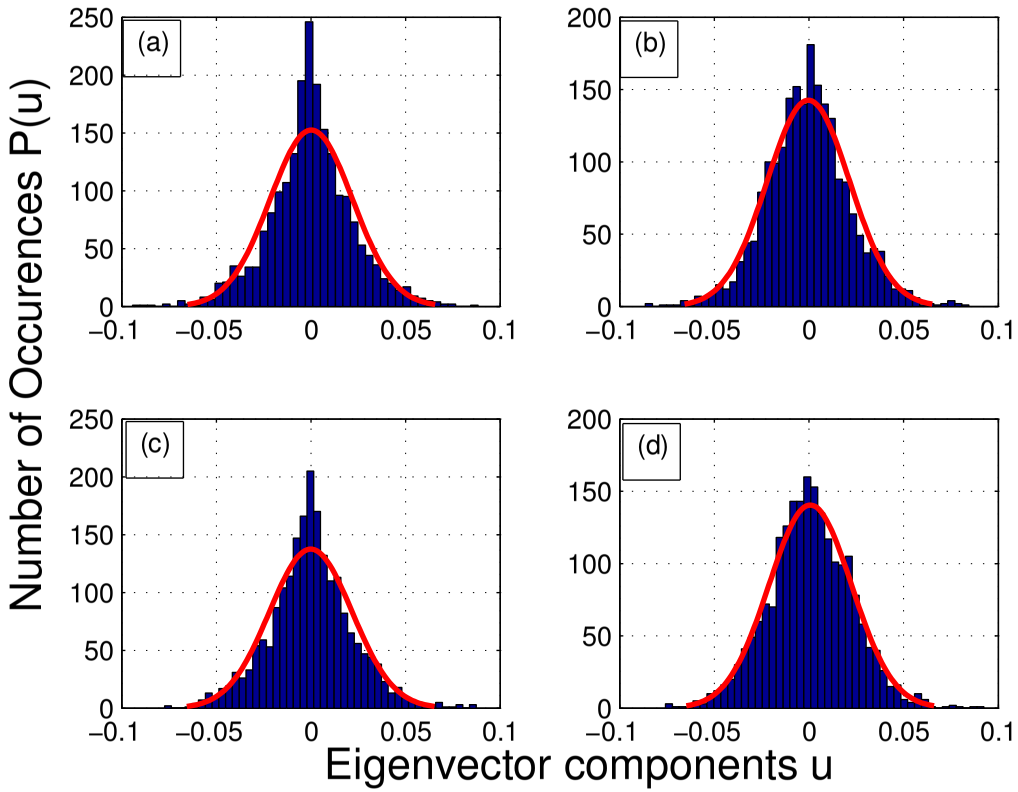

Figure 4.3: Comparison of Eigenvector Components of the AUTHORTYPICAL data set (D_2), (a) largest Eigenvector, (b) second largest Eigenvector, (c) third largest Eigenvector and (d) Eigenvector from the bulk

In order to remove the effects of the largest eigenvalue we use the techniques described in Section 4.2.2. We remove the contribution of $G^{large}(t)$ to each time series $G_i(t)$, and construct $\mathbf{C}$ from the residuals $\epsilon_i(t)$ of Equation (4.4). Fig. 4.4 for the AUTHORTYPICAL data set D_2 shows that the distribution $P(C_{ij})$ thus obtained has a smaller average value $<C_{ij}>$, showing that some of the Cross-Correlation contained in $\mathbf{C}$ can be attributed to the influence of the largest eigenvalue and its corresponding eigenvector.

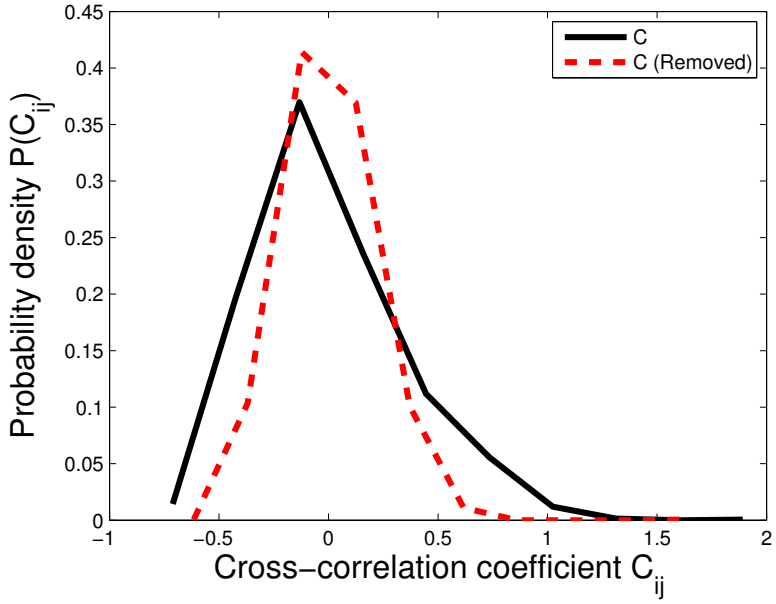

Figure 4.4: Probability distribution P of the Cross-Correlation coefficients of the AUTHORTYPICAL data set (D_2) for data before (black) and after (red) removing the effect of the largest eigenvalue by linear regression method

Inverse Participation Ratio (IPR): So we next focus on the remaining eigenvalues to see whether these relate also to key sources or major events and what information these contribute additionally to the images. The IPR quantifies the reciprocal of the number of eigenvector components that con-

tribute significantly. Fig. 4.5 (a) for the AUTHORTYPICAL data set D_2 shows I^k for the case of the control of Equation (4.1). The average value of I^k is $<I>\approx 0.0014 \approx 1/N$ with a very narrow spread, indicating that the vectors are "extended" (Lee & Ramakrishnan 1985) - i.e., almost all component elements (of the vector) contribute, with fluctuations around this average value confined to a narrow range. Fig. 4.5 (b) shows I^k for the Cross-Correlation Matrix constructed from the 2096 SenseCam images. The edges of the eigenvalue spectrum of $\mathbf{C}$ show significant deviations of I^k from $<I>$, indicating that there are major events contributing also to these eigenvectors. In addition, we find a number of small eigenvalues deviating from the control case, which suggests that these vectors are localized (Lee & Ramakrishnan 1985) - i.e. only a few images contribute to them.

Component content: Examination of the eigenvalue and eigenvector content indicates that the largest percentage of noise ($\approx 20\%$) for that period is described by the wearer 'working in front of the laptop' for a long time without performing any other activities. The eigenvalues deviating from the RMT upper bound indicate the wearer 'moving from in front of the laptop and preparing to go home', with every image capturing different moments of this event. For example, considerable light is captured at the moment of standing up, different colours appear when she turns around, etc. Although the images are visually very diverse, all have been captured in the same space, i.e. in the office. The eigenvalues deviating from the RMT lower bound involve several that are implicated in different minor activities; events such as 'commuting from home to the work place', the wearer 'talking with her colleague', 'the wearer sharing lunch with her colleague' etc. Only a few images, at a time, de-

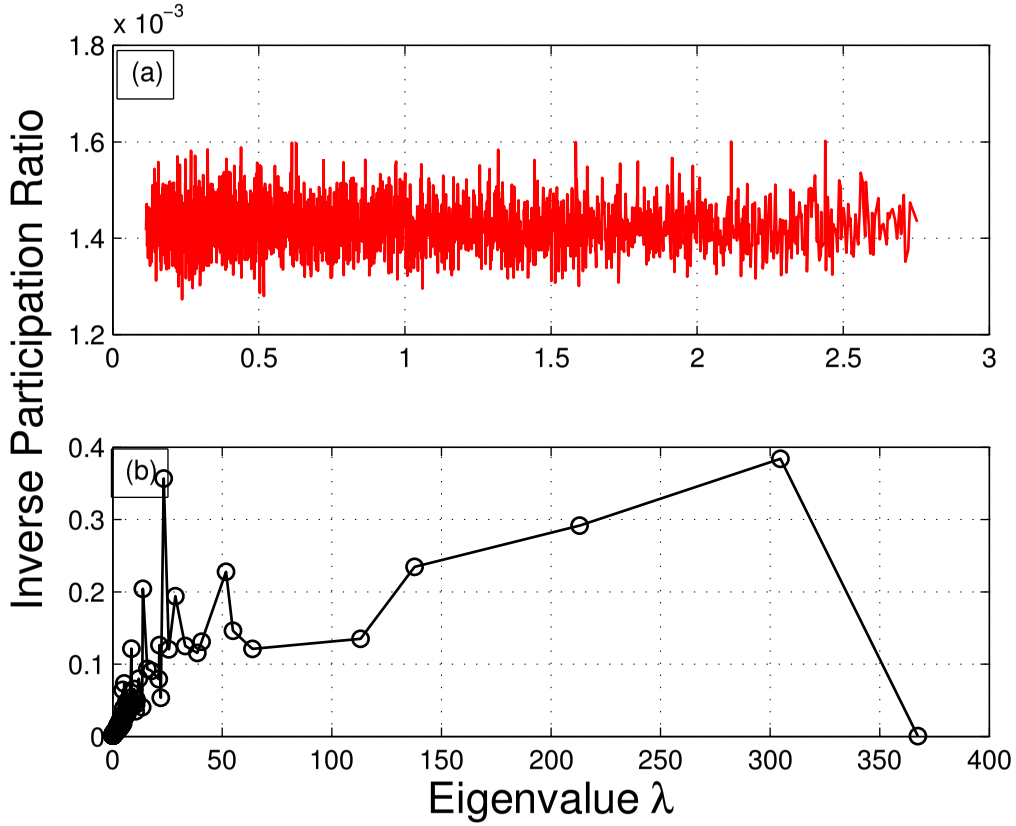

Figure 4.5: Inverse Participation Ratio (IPR) of the AUTHORTYPICAL data set (D_2) as function of eigenvalue λ for the (a) Random Matrix R and (b) Cross-Correlation Matrix C

pict the same 'environment'. We argue that this confirms our observation that the IPR shows that smallest eigenvalue deviations are 'localized', i.e. 'capture short-lived scenes'.

The distribution of the components of the remaining largest eigenvectors and their deviations show some distinctive clustering in Fig. 4.6 (a). In particular, Events {2, 6, 12}, Events {4, 5} and Events {3, 7} (in Table 2.2) are the major contributors here. Examination of the images indicates that each cluster reflects quite similar light levels for a group of events. For instance in the cluster of Events 2, 6 & 12: Event 2 describes the user 'passing from outdoors into the office'; The office is dark; When the user switches on the lights, light level changes are immediately captured. Similarly for Events {6, 12}, when the user is 'walking through the building', the camera sometimes 'registers' the lights and sometimes not. When the user 'prepares to leave the office', i.e. packing and, finally, standing, the camera captures lights from the ceiling, several images of the (dark) bag and so on. This light blockage strongly affects the information content from the sequence of images. Events {3, 7} are also similar: the scenario is that of 'the user sitting in front of her laptop, with laptop, lights and seating position unchanged over on extended period', contributing similar pixel values to this sequence of images. However, Event 11, although apparently describing a similar scenario (working at the PC), has not been grouped into the same cluster. Re-examination of these images indicates important differences in the scenarios: while the working period was long, the camera was inadvertently blocked by the user on numerous occasions. Clusters also emerge for other deviating eigenvalues, shown in Fig. 4.6 (b) and (c). However, events {2, 5, 6, 8, 9, 12}, {1, 3, 10}, {1, 2, 5, 8, 12}

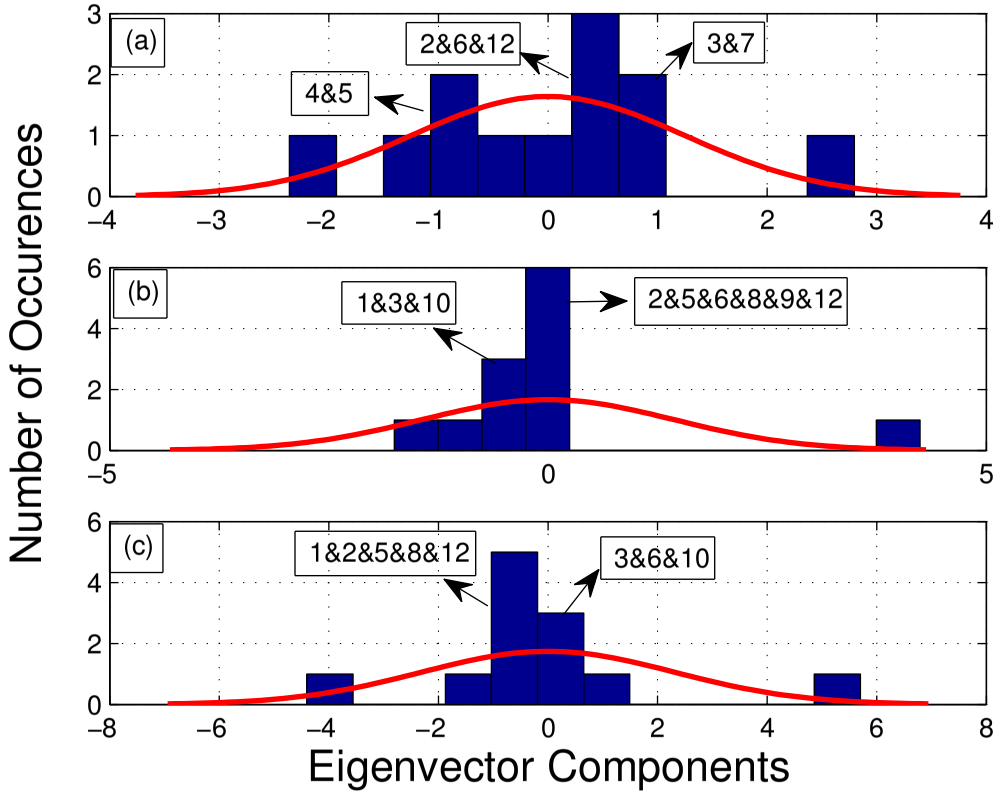

Figure 4.6: Distribution of eigenvector components: (a) largest remaining eigenvalue (b) 2nd largest and (c) 3rd largest remaining eigenvalue for the AUTHORTYPICAL data set (D_2)

and {3, 6, 10} mainly contribute to the second and third largest remaining[27] deviating eigenvector, respectively. This implies that these eigenvectors may carry additional information on the description of possible lead-in, lead out to major event since the sets of events share common features.

4.4 Summary

In this chapter, we studied whether Random Matrix Theory (RMT) can be applied to extract meaningful information and noise from a data sets. Significant deviations of the signal eigenvector from those of from RMT predictions are observed. Further, we analyze the deviations from RMT, and find that (a) different eigenvectors encapsulate different features, suggesting that different major events are captured by various eigenvectors. (b) By examining the eigenvectors corresponding to identifiable image components, we find that alternating light levels as a key source of deviation are always picked up by the device in terms of contributions to the largest remaining deviating eigenvector. In addition, similar events have been grouped together, and captured by distinctive clusters in the eigenvector distribution. (c) The second and third largest remaining deviating eigenvectors have a similarly-clustered distribution to that of the largest remaining deviating eigenvector. However, the different clusters here relate to different events. We take from that these eigenvectors carry additional information to the description of events leading-in or leading-out to major events. (d) We also note that even 'quite similar' scenarios, stemming from identifiably slightly different key sources[28], are distin-

[27]The smaller eigenvalues, and their corresponding eigenvectors

[28]Change of illumination levels: from outdoor to indoor, block or unblock the computer screen, switch on/switch off the lights;

guishable; these events have been classified by different clustered distributions. This implies such key sources play a major part in classification by the Cross-Correlation Matrix. The proposed cleaning technique of separating the noisy part from the non-noisy part proved clearly useful. Overall, the RMT technique provides a powerful tool in analysis of cross-correlations across whole data streams.

Chapter 5

Finding Motifs in Large Personal Lifelogs

5.1 Introduction

In previous chapters, we introduced and evaluated the use of sophisticated time series analysis methods for interpretation of large lifelogging SenseCam data sets, generated by the author. The results of this evaluation suggest that strong correlations do exist in lifelogging image time series, with recognisable cycles representing specific events. In this chapter, we build on this observation to address the challenge of refining the analysis of lifelogs by studying time series motifs. Such motifs are frequently occurring, (but often previously unknown), subsequences of longer time series (Lin et al. 2002). Many researchers have studied the extraction of characteristics features from multi-dimensional time series data (Nevill-Manning et al. 1998, Androulakis 2005, Minnen et al. 2006). Given that visual lifelogs contain records of a wearer's daily activities and events that occur over different time periods, we argue that motifs can be used

to identify and represent activities of different length and timing in such data (which share common features with big data types). We explore this idea by analysing high frequency patterns in multi-dimensional visual lifelogging data.

5.2 Methods

In Chapter 3, we studied the spectral dynamics of SenseCam images by applying the multi-scaled Cross-Correlation Matrix technique, for time series exhibiting atypical or non-stationary characteristics, symptomatic of 'Distinct Significant Events' in the data. This suggests that we can use such *key episodes* to identify boundaries between different distinct daily events. The results indicated that such events or activities can be detected at different scales through wavelet analysis. Building on this observation, we aim in this chapter to extract motifs or extended sub-sequences at different wavelet scales using the Minimum Description Length (MDL) principle. Our method is first to calculate the Cross-Correlation Matrix structure and the Maximum Overlap Discrete Wavelet Transform (MODWT) (introduced in Chapter 3, sections 3.2.2 and 3.2.3). We then employ the Symbolic Aggregate approXimation (SAX) algorithm for discretization of time series data into symbolic strings. Finally we detail our motif extraction algorithm, based on the MDL principle.

5.2.1 Symbolic Aggregate approXimation (SAX)

Time series are always complex and large, in order to reduce the magnitude of time series, some of dimensionality reductions methods are used, they are including: DFT (Discrete Fourier Transform), DWT (Discrete Wavelet Trans-

form), PAA (Piecewise aggregate Approximation). In this chapter, we use dimensionality-reduction algorithm called Symbolic Aggregate approximation (SAX) based on PAA (Lin et al. 2002, Keogh et al. 2001, Lin et al. 2007). PAA is dividing a time-series into several segments and calculating the average value in each segment. We apply this technique to transform the largest eigenvalue time series into a sequence of symbols. For the largest eigenvalue time series $\lambda_1(t)$ (obtained from Equation (3.5)) with number of images n (n is length of the time series), then time series n can be reduced to a string of arbitrary length w, (where $w < n$, typically $w << n$) and an alphabet size of arbitrary length a, (where $2 <= a <= 10$). In SAX, parameters w and a setting are depending on the data set. In this thesis, we implemented SAX algorithm in MATLAB with local machine (8GB RAM), we tried parameter w from 8 to 16 and parameter from 3 to 5, the results obtained are quite similar. The largest eigenvalue time series $\lambda_1(t) = \{x_1, ..., x_n\}$ of length n can be represented as a w-dimensional space by a vector $\bar{\lambda} = \{\bar{x}_1, ..., \bar{x}_w\}$:

$$\bar{x}_i = \frac{w}{n} \sum_{j=\frac{w}{n}(i-1)+1}^{\frac{w}{n}i} x_i \tag{5.1}$$

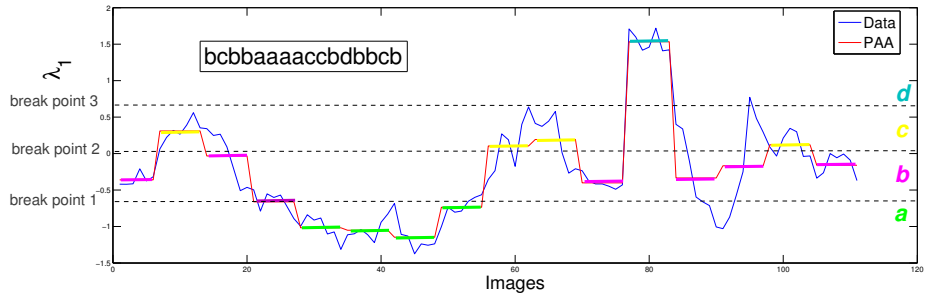

Figure 5.1: Example of time series transformation into SAX symbols. Here, n=112, w=16, a=4. The time series is mapped to the PAA symbols **bcbbaaaaccbdbbcb.**

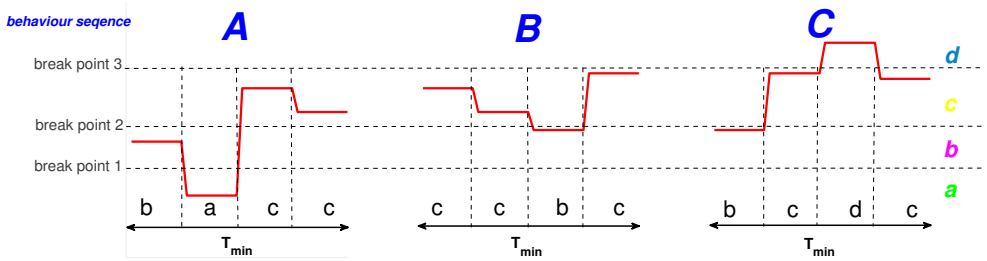

Figure 5.2: Example of 'behaviour symbol' assignment for pattern order of PAA symbols. In this, **A=bacc, B=ccbc and C=bcdc.**

In order to transform the vector of w dimension into a sequence of 'PAA symbols', it is necessary also to determine 'breakpoints' that determine the range of the PAA value for assigning unique PAA symbols. One approach is to determine the *breakpoints* that will produce an equal-sized area under a Normal distribution. *Breakpoints* are a sorted list of numbers $B = \{\beta_1, ..., \beta_{a-1}\}$ such that the area under a $N(0,1)$ Standard Normal distribution from β_i to $\beta_{i+1} = 1/a(\beta_0)$ and β_a are defined as $-\infty$ and ∞, respectively. Once the breakpoints have been obtained we can discretize a time series as follows. We first obtain a PAA of the time series; All PAA coefficients that are below the

smallest breakpoint are mapped to the symbol 'a', all coefficients greater than or equal to the smallest breakpoint and less than the second smallest breakpoint are mapped to the symbol 'b', etc. Fig. 5.1 illustrates the idea; Finally, a 'behaviour symbol' is assigned to every subsequence of PAA symbols. An example is given in Fig. 5.2, for which the analysis window of T_{min} is the minimum length of motif for the time series.

5.2.2 Estimating Extracted Motif Candidate Based on MDL Principle

Several theoretical information theory principles from literature are relevant to the current analysis, including AIC (Akaike's Information Criterion), BIC (Bayesian Information Criterion) and MDL (Minimum Description Length) principles (Pitt & Myung 2002).

The AIC estimates the best model based on 'prediction capability', while BIC estimates the best model based on Bayesian principles. Our approach is focused however, on finding frequent patterns, rather than prediction, for the time series. The *MDL principle* states that the best model to describe a set of data is that which minimises the description length of the *entire* data set. The underlying concept is the selection of the best model to compress the data. Table 5.1 summarises the principal notation used in this sub-section.

The 'data encoding cost' is the lower bound of description length that is required to encode each segment. The 'parameter encoding cost' is the description length that is required to describe the order of the behaviour symbol (BS) in each segment. Finally, the 'segmentation cost' is required to describe the location of all segments. The work-flow of the MDL pattern algorithm

Table 5.1: Summary of the Notation of the SAX Algorithm

Symbol	Meaning
$\lambda_1(t)$	A time series $\lambda_1(t) = \{x_1, ..., x_n\}$
$\bar{\lambda}$	A Piecewise Aggregate Approximation of a time series $\bar{\lambda} = \{\bar{x}_1, ..., \bar{x}_w\}$
$\hat{\lambda}$	A symbol representation of a time series $\hat{\lambda} = \{\hat{x}_1, ..., \hat{x}_w\}$
w	The number of PAA segments representing time series $\tilde{C}$
a	Alphabet size(e.g., for the alphabet=a,b,c,d, a=4)
T_{min}	Analysis window (e.g., for the alphabet A=bacc, T_{min}=4)
BS	Behaviour Symbol (e.g., for the alphabet A=bacc)
BSS	Behaviour Symbol Sequences (e.g., ACDFHCBDFHDADFHBED....in Fig. 5.3 (a))
DL	Description Length

can be visualized in Fig. 5.2. For example, in this figure, the length of the first segment is s_1=7, the length of the second segment is s_2=3 and so on. In addition, we assume that the j_{th} BS has a length l_{ij}. A data encoding cost for the j_{th} BS in the i_{th} segment is calculated then as:

$$- l_{ij} log_2 \frac{l_{ij}}{t_i} \tag{5.2}$$

By calculating this cost for all unique Behaviour Symbol Sequences (BSS) in the i-th segment, we obtain the data encoding cost of the whole segment as:

$$\sum_j -l_{ij} log_2 \frac{l_{ij}}{t_i} \tag{5.3}$$

Using the following equation, we then calculate the data encoding cost $DL1(\tilde{C}|SC)$ of $\tilde{C}$ that is segmented by the pattern SC:

$$DL1(\tilde{C}|SC) = \sum_i^m \sum_j -l_{ij} log_2 \frac{l_{ij}}{t_i} \tag{5.4}$$

We calculate the complementary parameter encoding cost of each segment as $log_2 t_i$. Then, the second segment cost $DL2(\tilde{C}|SC)$ of $\tilde{C}$ is calculated as:

$$DL2(\tilde{C}|SC) = \sum_i^m log_2 t_i \tag{5.5}$$

and the segmentation cost $DL3(\tilde{C}|SC)$ as:

$$DL3(\tilde{C}|SC) = m log_2 (\sum_i^m t_i) \tag{5.6}$$

Finally, based on this table, we obtain the description length of $\tilde{C}$ that is segmented by the pattern SC as follows:

$$MDL(\tilde{C}|SC) = DL1(\tilde{C}|SC) + DL2(\tilde{C}|SC) + DL3(\tilde{C}|SC) \tag{5.7}$$

We use Equation (5.7) as the MDL estimation function for the MDL pattern detection algorithm.

(a) a BS sequence:

	A	C	D	F	H	C	B	D	F	H	D	A	D	F	H	B	E	D		
(BS length)	2	5	1	1	1	5	3	1	1	1	1	2	1	1	2	2	3	1		

(b)

	A	C	D	F	H	C	B	D	F	H	D	A	D	F	H	B	E	D
	2	5	1	1	1	5	3	1	1	1	1	2	1	1	2	2	3	1
Segment	s1		s2			s3		s4			s5		s6			s7		

(c)

Segment	Length	Data encoding cost	Parameter encoding cost
s1	7	A x 2 : $-2\log_2 \frac{2}{7}$ C x 5 : $-2\log_2 \frac{5}{7}$	$\log_2 7$
s2	3	D x 1 : $-1\log_2 \frac{1}{3}$ F x 1 : $-1\log_2 \frac{1}{3}$ H x 1 : $-1\log_2 \frac{1}{3}$	$\log_2 3$
s3	8	C x5 : $-5\log_2 \frac{5}{8}$ B x3 : $-3\log_2 \frac{3}{8}$	$\log_2 8$
		. . .	
s7	6	B x 2 : $-2\log_2 \frac{2}{6}$ E x 3 : $-3\log_2 \frac{3}{6}$ D x 1 : $-1\log_2 \frac{1}{6}$	$\log_2 6$
Sum	**34**	**DL1**	**DL2**

And with Segmentation cost: **DL3 = 7 log 34**

(d) MDL= DL1 + DL2 + DL3

Figure 5.3: Calculation of the MDL pattern algorithm

5.3 Results

In order to evaluate our technique, we tested our method on two data sets: the AIHSWEDS data set (D_3)[29] and the NTCIR-12 (3 USERS) data set (D_4)[30] have been introduced above in Chapter 2 Section 2.6.

First, the MODWT of the pixels for each image was calculated within each window of size 400 images and the Cross-Correlation Matrix between pixels at each scale found. The eigenvalues of the Cross-Correlation Matrix in each window were determined, and the eigenvalue time series was normalised in time. Then, the largest eigenvalue for different window sizes was calculated. Finally, the SAX algorithm was applied to transform the time series to PAA symbols.

Fig. 5.4. illustrated the AIHSWEDS data set (D_3), where wavelet scales 1-4 correspond to a 1-2 minute period, a 2-4 minute period, a 4-8 minute period and a 8-16 minute period, respectively. The different features, found at various scales, suggest that the Cross-Correlation Matrix captured different major events with different time horizons. The largest Eigenvalue dynamics show that with increased wavelet scales comes increased smoothing, as expected. This removes some of the high frequency small-scale changes, typically associated with noise.

[29]From the AIHS data set includes 7,549 images from one lifelogger
[30]From the NTCIR-12 Lifelog data set includes 34,758 images from three lifeloggers

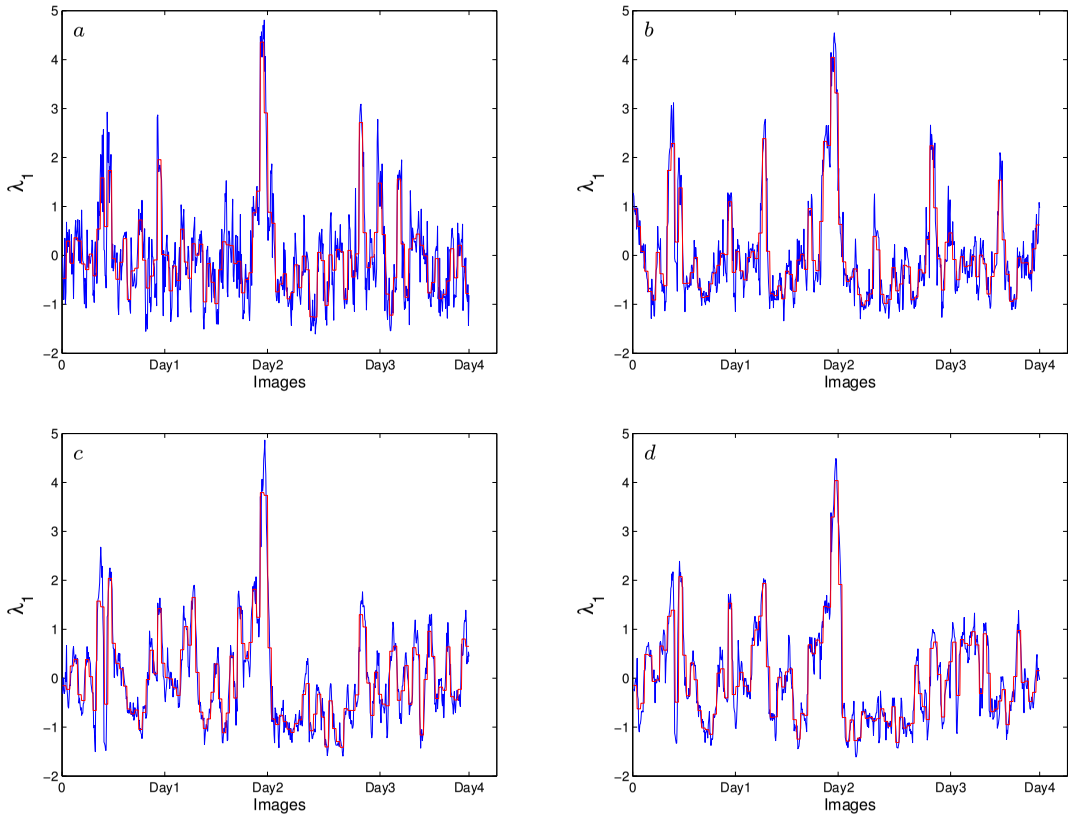

Figure 5.4: The largest eigenvalue λ_1 (*blue*) dynamics with the curve generated for the PAA (*red*). (a) Wavelet scale 1, (b) scale 2, (c) scale 3 and (d) scale 4 for the AIHSWEDS data set (D_3).

Fig. 5.5 depicts examples of motifs discovered from different wavelet scales using the method we introduced in Section 5.2 for the AIHSWEDS data set (D_3). As depicted, different motifs were extracted from different wavelet scales. Since different wavelet scales correspond to different time horizons, the motifs, extracted from different scales, should represent different 'events' that the wearer experienced each Wednesday.

Examining the images that are identified by the motif analysis, we find that both a_1 and a_2 describe the combined event of driving back home in the afternoon, followed by watching TV. Motif b_1 combines activities of eating, and then moving to the living room. Motif b_2 corresponds to driving to the shopping mall. Both c_1 and c_2 are similar events where the wearer drives to work and then sits for some time in front of the computer. For d_1 the event sequence comprises sitting and watching TV, talking with family and then starting to cook, while d_2 comprises sitting in the shopping mall and then driving to an outdoor garden with children.

By examining the data set, we determine that most motifs discovered at the same wavelet scale represent similar scenarios. We note also that most describe the change from a static setting (e.g., sitting in front of a computer or TV) to a more dynamic activity. Light change remains an important feature in identifying/distinguishing between *key episodes* detected by our technique. An added value here is a way of aggregating related or sequential events in a recognisable template, which can be used as potential marker for automatic extraction or comparison of similar episodes from an extended time series.

The descriptive statistics for NTCIR-12 (3 USERS) data set (D_4) is reported in Table 2.5. The organisers of the NTCIR-12 Lifelog also identified

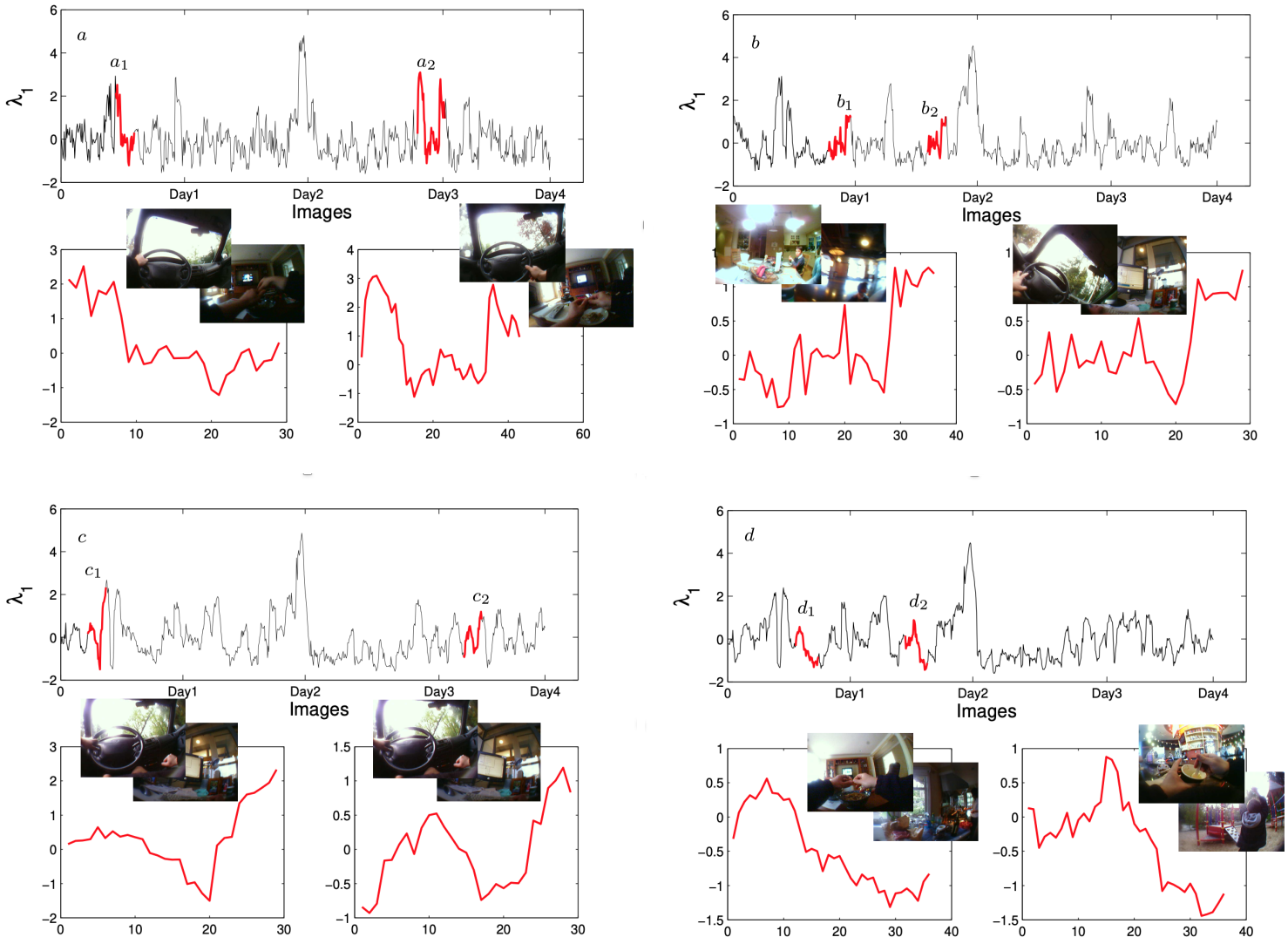

Figure 5.5: Example of Motifs extracted based on MDL principle from different wavelet scales. (a) Wavelet scale 1, (b) Wavelet scale 2, (c) Wavelet scale 3 and (d) Wavelet scale 4 for the AIHSWEDS data set (D_3).

about 1,000 concepts by using the CAFFE visual concept detector. The output of this concept detector, however, cannot be used to evaluate our method since it includes recognition only of simple visual objects such as a desktop computer, beer glass, banana, car wheel and so on. Moreover, the annotations provided are not very accurate. In order to evaluate our methodology, therefore, we first had to create a gold standard by identifying key lifelogger activities. We manually annotated activities that the three lifeloggers performed over ten days, resulting in annotations for a total of 34,758 images. Descriptive statistics of these extracted data are reported in Table 5.2.

Fig. 5.6 shows the time series of the largest Eigenvalue dynamics across different wavelet scales for the NTCIR-12 (3 USERS) data set (D_4) from three lifeloggers. The heatmaps show clearly different life patterns for the three. For example, for lifelogger 1, some areas are consistently captured by the camera at certain scales, such as the section of map around Day 8, (captured by wavelet scales 4, 5, 6 and corresponding to periods of 8-16, 16-32 and a 32-64 minutes, respectively). These peaks refer to periods when the lifelogger was sitting in the living room; the camera consequently captures the ceiling lights. For lifelogger 3, the section of the map around Day 1 at high frequency scales and section of the map around Day 7 at middle frequency scales are consistently captured by the camera. This represents light captured from similar appliances - as in the previous case. By examining the data sets, we note however that lifelogger 3's lifestyle is more active and veried compared to the other two lifeloggers. Activities include passing through airport and train stations, visiting the pub, attending a party in a friend's home and so on. The Fig. 5.6 heatmap reflects the increased number of peaks in lifelogger 3's time series compared to the

Table 5.2: Summary of Ground Truth of the NTCIR-12 (3 USERS) data set (D_4)

Lifelogger	**Events**	**No.**	**Lifelogger**	**Events**	**No.**	**Lifelogger**	**Events**	**No.**
1	Attending Lecture	2	2	Attending Lecture	1	3	Cooking	12
1	Cooking	35	2	Cooking	17	3	Drinking	33
1	Doing Laundry	3	2	Cycling	7	3	Driving	13
1	Drinking	43	2	Doing exercise	3	3	Eating	23
1	Driving	37	2	Doing Laundry	3	3	Ordering Food	10
1	Eating	34	2	Drinking	52	3	Playing Lotto	2
1	Fixing Car	3	2	Eating	27	3	Playing with Phone	99
1	Giving Lecture	5	2	Ordering Food	15	3	Reading paper	20
1	Ordering food	20	2	Playing Guitar	6	3	Shopping	5
1	Playing with Phone	158	2	Playing with Phone	86	3	Sitting	29
1	Reading Paper	37	2	Reading Paper	29	3	Standing	32
1	Shopping	3	2	Shopping	4	3	Talking with People	89
1	Sitting	228	2	Sitting	135	3	Using Coffee Machine	2
1	Standing	107	2	Standing	68	3	Walking	130
1	Talking with People	162	2	Talking with People	99	3	Watching TV	47
1	Using Coffee Machine	5	2	Using Coffee Machine	1	3	Working in front of Computer	55
1	Walking	236	2	Walking	170			
1	Watching TV	80	2	Watching TV	6			
1	Working in front of Computer	163	2	Working in front of Computer	114			

other two. The heatmap shows that some features such as watching TV and working in front of computer are consistently captured at certain scales, while others, such as light changes, features across scales, suggesting that the Cross-Correlation Matrix successfully highlight different major events with different time horizons.

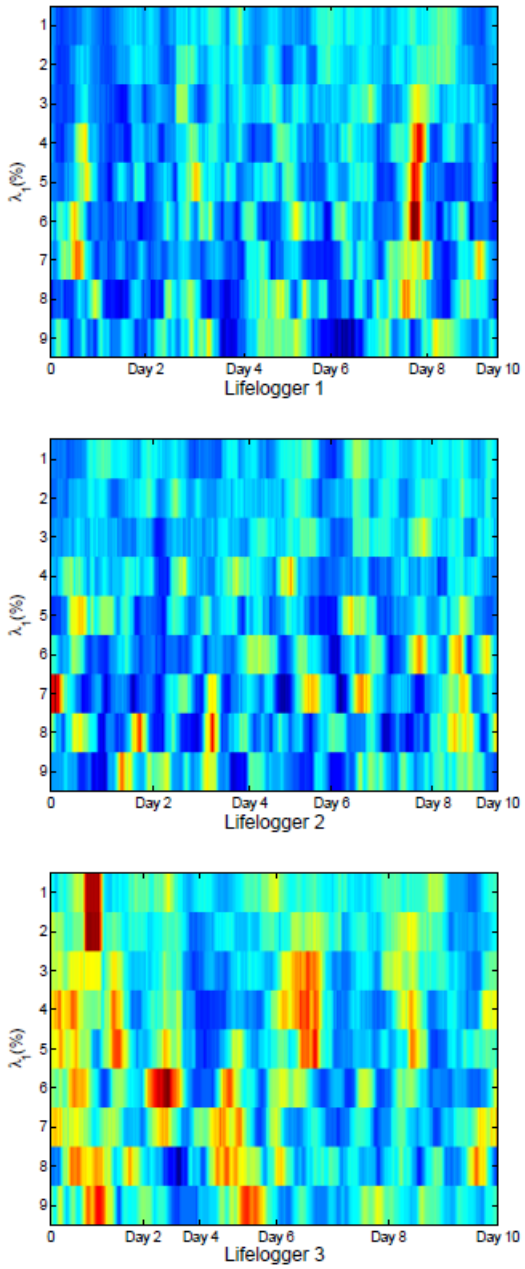

Figure 5.6: Heatmap diagram showing the dynamics of the largest Eigenvalue λ_1 across 9 wavelet scales. Scales 1 to 9 correspond respectively to periods of 1-2 , 2-4, 4-8, 8-16, 16-32, 32-64, 64-128, 128-256 and 256-512 minutes, respectively, for the NTCIR-12 (3 USERS) data set (D_4).

As shown in Table 5.3, our approach extracted key 'motifs' such as 'working in front of a computer'. We defined a 'matching' pair as similar event 'identification', while a 'mismatch' was recorded if the event occurred but was not identified by the motif technique (i.e. a false negative). The results suggest that the high frequency wavelet scales perform better for lifeloggers 1 and 3 while the middle frequency wavelet scales are more accurate for lifelogger 2. Unfortunately, the number of 'match' pairs is less than that of 'mismatch' pairs for each wavelet scale. The main weakness, as well as strength, is that different wavelet scales highlight different distinct events dependent on the time horizons. In our example, working in front of the computer can last for several hours or a few seconds, so that some additional measure indicating *event duration* is required. In the wavelet approach also, some events are missing from some scales but prominent in others. We found that a few extracted 'motifs' in consecutive wavelet scales do overlap, so that some aggregation is possible to address in part of this issue. We can roughly calculate 'our approach to identification' accuracy by adding the 'match' pairs of wavelet scales 1, 3, 5, 7 & 9 and dividing by the total number of events of identified type. Identification accuracy for lifeloggers 1, 2 & 3 is $\sim$ 40%, 76% and 65% of total events, respectively. We note that while this accuracy is not ideal and suggests that further modification is needed for motifs consecutively repeated or persistent over time, this does seem to indicate that frequent activitity change may improve identification and help in refining the wavelet application.

Table 5.4 shows the accuracy of different motifs cross different wavelet scales. In general, the middle wavelets 3 to 5 (corresponding to 4-8, 8-16 and16-32) give us the best results on the motifs "Playing with Phone", "Talk-

Table 5.3: Motif: 'Working in front of Computer' Event

Wavelet Scales	**User 1**		**User 2**		**User 3**	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
W1	9	154	7	107	5	50
W2	13	150	8	106	11	44
W3	17	146	11	103	7	48
W4	24	139	12	102	15	40
W5	27	136	22	92	19	36
W6	13	150	37	77	11	44
W7	9	154	43	69	5	50
W8	3	160	20	94	0	55
W9	3	160	4	110	0	55

Table 5.4: The accuracy of different motifs

Motifs	**User 1**		**User 2**		**User 3**	
	Best Wavelet Scales	Accuracy	Best Wavelet Scales	Accuracy	Best Wavelet Scales	Accuracy
Playing with Phone	W4	10.12%	W3	18.60%	W3	15.15%
Talking with People	W2	10.49%	W4	15.15%	W4	24.71%
Walking	W5	15.38%	W4	13.53%	W5	14.62%
Working in front of Computer	W5	16.56%	W7	37.72%	W5	34.55%

ing with People", "Walking" and "Working in front of Computer". It was noted that the accuracy of results vary from 37.72% to 10.12%. This is due to two main reasons: First, our method is powerful in extracting activity changes, such as light level changes, so the single activity evaluation depends on before and after activities; Second, as we mentioned in Chapter 3, the main weakness as well as strength for wavelet scales is that different scales captured different events depending on the time horizons. Some events are missing by certain wavelet scales. It was also noted on the same motif, for example "Working in front of Computer", different users received the best results from different wavelet scales. It indicates the results linked with different user profiles (lifestyles). One of the promising approaches is transforming the visual Lifelogging images into semantic contexts. The semantics contexts will help us to better understand visual Lifelogs images.

5.4 Summary

The major contributions of this chapter include exploration of MODWT & motif approaches for investigation of two well-known lifelogging data sets[31]. Our methods demonstrate considerable potential for the organisation, structuring and interpretation of vast amount of heterogeneous streams of visual data. In particular, through application of the Maximum Overlap Discrete Wavelet Transform (MODWT) on equal-time Cross-Correlation Matrix, we find that different features occur at different wavelet time-scales. This suggests that the Cross-Correlation Matrix captures different major events corresponding to different *time* horizons. Further, the discovery of distinct behavioural motifs

[31]AIHS & NTCIR-12(subsets)

provides a basis for prototype templates for identification of similar scenarios at specific time scales e.g., 'typical' lifestyle patterns of lifeloggers. Indications are that wavelet methods may be more effective for identification of motifs in time series with frequent activity changes, but clearly this requires further evaluation. For less active lifestyles, it seems likely that additional statistical information criteria will also need to be used in enhancement of the technique. Nevertheless, the identification of specific behavioural motifs does indicate potential for refinement of the method.

Chapter 6

Conclusions and Future Work

6.1 Summary

Visual Lifelogging is used to describe the process of tracking personal activities by using wearable cameras. A wearable camera such as the SenseCam generates large volumes of continously streamed personal data. The key challenge is in extracting meaningful information from such large data sets in order to make these useful for the user. Although deep learning techniques for improved classification of images have increased in popularity, lifelogging presents difficulties, due to the lack of annotation, incompatible lifelogging devices and incompleteness of data sets; (errors occurring during the lifelogging process). In consequence, promising results for lifelogging image data continue to prove elusive.

In this thesis, for the first time, we considered lifelogging image data as a Complex System, (with typical shared features). We proposed a novel approach based on application of time series methods, commonly used for analysis of various other Complex Systems. Several large lifelogging image data sets

were generated by different users, and from different lifelogging devices, and were used to test the approach. The time series analysis methods aimed to restructure the lifelogging data sets, identify common causes (or factors) of the dynamics, and to identify and extract common life 'patterns' from the lifelogs. In order to achieve these aims, we initially applied the Detrended Fluctuation Analysis (DFA) method to the SenseCam image time series. The results show that ((i) a) the data set is not a random walk and ((i) b) strong correlation exists in the time series with recognisable cyclic fluctuation. The DFA method provided the initial background information for SenseCam.

We further applied the equal-time Cross-Correlation Matrix method to characterise *dynamical changes* in non-stationary multivariate SenseCam time series. The behaviour of the largest eigenvalue of a Cross-Correlation Matrix with different sliding window sizes was studied. Results show that ((ii) a) large window size gives smoother series, as expected, and highlights major features with some peaks in the series becoming more pronounced. The sliding window approach thus helped to remove some of the high frequency small-scale changes, many of which can be treated as noise. Results also suggest that ((ii) b) the largest eigenvalue can be used to explore information from the image series, which reflects the largest change overall. This was found to apply for data from several user SenseCam recordings in different data sets. By applying the Maximum Overlap Discrete Wavelet Transform (MODWT) and from the information present in the largest eigenvalue, we further found that ((ii) c) different features occur at different wavelet time-scales suggesting that the Cross-Correlation Matrix of different major events corresponds to different time horizons. In addition, by examining the change in eigenvalue-ratios over

time, we confirmed that the largest eigenvalue carries most of the major event information (as for (ii) b), while subsequent eigenvalues carry information on supporting or lead-in/lead-out events ((ii) d). These consistently occurring peaks thus help to identify major and subsidiary events of lifelogging data. For the major events, it is clear that light level is a major delineator during static periods of image sequence ((ii) e).

We also considered the effects of noise on the image time series. Random Matrix Theory (RMT) was exploited to filter noise from the Cross-Correlation Matrix, constructed using lifelog data streams. Significant deviations from RMT predictions were observed. In analysing these deviations confirmation was obtained ((iii) a) that the largest eigenvalue and its corresponding eigenvector present information from the image that reflects the largest change in the image data recording, and ((iii) b & c) the smallest eigenvalues, and their corresponding eigenvectors, represent *short duration major events* from the SenseCam recording. The 'cleaning technique' (of separating the noisy part from the non-noisy part) is demonstrated to be particularly useful ((iii) d).

We also proposed using *motif discovery techniques* based on SAX discretization and wavelet methods to investigate subsets of two well known Lifelogging data sets, AIHSWEDS data set (D_3) and the NTCIR-12 (3 USERS) data set (D_4). We showed that ((iv) a) Minimum Description Length (MDL) and wavelet analysis can be used to extract motifs (repeated patterns of different length) from multi-dimensional time series data. The motifs discovered ((iv) b) *provide prototype templates for identification* of similar scenario behaviours at specific time scales e.g. 'typical' lifestyle patterns of lifeloggers such as that of 'working at the computer'. The evidence from motif methods pro-

vides additional confirmation of the previous findings on the nature of major events, indicating that light changes are natural markers for distinguishing key episodes (as (ii) e). Further, ((iv) c) while behavioural motif identification is modest for some data sets, it is more successful in the case of extended and multiple activities, suggesting potential for refinement by means of statistical information criteria and weighting of events by type.

6.2 Future Work

The ideas and methods presented in this thesis provide new insights on how to structure, organise and analyse lifelogging images. Given the novelty of the approach described, there are a number of research challenges that can be undertaken in the future.

From the data perspective: although the concept of lifelogging is not that recent, large lifelogging data sets of decent quality are still very limited. Such data contain personal information and consequently, privacy and security issues are a concern for most lifeloggers (Li & Hopfgartner 2016). In this thesis, where we have examined two public domain data sets, referred to as 'AIHS' and 'NTCIR-12', these are far from perfect, due to incompatible recording errors caused by human error and devices limitations or differences. For example, the AIHS data set reportedly includes 19 days with a total of 45,612 images, but if each image is examined, the actual coverage is 29 days with a total of 43,399 images. Similarly, for the NTCIR-12 Lifelog data set; e.g., for Lifelogger 2 on the tenth day of collection, some images are not in the correct order while on the ninth day of Lifelogger 3, most images appear to be the same. In addition, (for there are many generations of lifelogging de-

vices available on the market), the quality of images produced by the different cameras vary significantly. By comparing images recorded by two lifelogging devices (SenseCam and Narrative), we notice that not only do camera lenses differ (fisheye vs normal)[32], but the SenseCam device captures images with a larger field of view (and of poorer quality due to movement and occlusion from the wearer's hands). In contrast, the images produced by the Narrative device in general have a normal field of view and are of much better quality than those of SenseCam. Clearly, it would be useful to evaluate the robustness of our methods by benchmarking on visual lifelogging data sets and devices with good enough resolution, low noise and accurate lifelogging recording. The ideal would be at least to ascertain what is 'good enough' to some not same extent defined by context.

Other future work could include investigation of the range of other sensors data inputs and their value towards, e.g., improved activity recognition. Accelerometer data could be an important addition in this respect, (see e.g., (Albatal et al. 2013)). Inference of contextual information about common daily activities such as sitting, walking, driving and lying down should be possible to detect. In particular, combination of accelerometer and image sensor data streams together may enable more accurate *event boundary identification* for lifelogging data. Distinguishing consecutive movements indicate that the associated image shift is quite likely to represent a boundary between different events or activities such as walking from home to work, walking from the office to lunch, walking from home to the shops, etc. but with different scope of environment, levels of personal interaction and so on. Combining sensor

[32]fisheye means wide-angle; normal means straight lines of perspective

information may thus help to classify lifelogging images into more meaningful activities. However, one immediate improvement in this regard might be to sample the different sensor data at different rates in order to apply the methods described. The most common among them are those using a fixed window size (Maurer et al. 2006) or sensor methods such as Kalman filter (Alatise & Hancke 2017) or convolutional neural network (Ordóñez & Roggen 2016).

From the implementation perspective: one limitation of the presented methods is that these require extensive computer memory and the speed to solve large systems of equations. In this thesis, all methods were implemented using MATLAB, which is widely accessible and regularly updated, but speed of computation is limited. In future work, a highly scalable High Performance Computing (HPC) framework should be considered, such as cluster or grid computing (Foster et al. 2008) or the recently extended HPC framework MapReduce/Hadoop[33] and Apache Spark[34]. Such a framework would help to evaluate larger lifelogging data sets using our methods, both to confirm initial results and help explore further the detailed features of lifelogging images. The availability of larger lifelogging data sets would enable more precise validation of the methods proposed, together with the potential to explore other features of these data in more detail, not least in order to provide typical real-time 'system alerts' for more immediate practical and personal feedback.

From the methods perspective: improvement of current motif discovery technique is a promising direction to explore. The SAX motif discovery algorithm employed in Chapter 5 has weaknesses in term of being *approximate*, *slow* and requiring *many parameters*, which have to be carefully chosen to en-

[33]https://hadoop.apache.org
[34]https://spark.apache.org

sure high precision. For future work, it would be worth while to consider the algorithm STAMP (Scalable Time series Anytime Matrix Profile) (Yeh et al. 2016) which promises *exact*, *parameter free* and *speed up* pattern identification. Further, it would be interesting to benchmark the results between the mSTAMP (Yeh et al. 2017) algorithm discovery of multidimensional motifs and STAMP with wavelet methods for similar type of lifelogging data.

In addition, in Chapter 5, we have used the dimensionality reduction of the symbolic representation approach for discovering repeated distinctive patterns (*motifs*) for the lifelogging images data. In the future, it would be useful to have a module to predict subsequent events given the context of an existing event. Recently, Hu (2016) shows that lifelog data (accelerometer sensors) can be modelled and predicted with suitable models and under several constraints. Model-based methods, such as Gaussian Mixture Models, Hidden Markov Models and autoregressive moving average (ARMA) models are popularly used for forecasting or predicting multivariate data sets (Williams et al. 1998, Krogh et al. 2001, Wiest et al. 2012, Eirola & Lendasse 2013, Li et al. 2018, Yang et al. 2018). However, for model comparison, measures of model fit and model complexity are required. Statistical analysis and mixed information criteria (such as Akaike's information criterion (AIC), corrected AIC (AICc), Bayesian information criterion (BIC), minimum description length (MDL), and the Hannan-Quinn (HQ)) provide a basis for interpretation of data, especially where these are incomplete and imperfect and offer a novel way to choose the "best" time series models (Yang 2005, Fonseca & Cardoso 2007, Dziak et al. 2017).

We provided some initial results from application of time series methods to

lifelogging image data, but there is still much work to be done. However, we believe that work presented in the this thesis provides a good starting point for future research work in this area.

Bibliography

Absil, P.-A., Sepulchre, R., Bilge, A. & Gérard, P. (1999), 'Nonlinear Analysis of Cardiac Rhythm Fluctuations using DFA Method', *Physica A: Statistical mechanics and its applications* **272**(1), 235–244.

Acharya, R., Lim, C. & Joseph, P. (2002), 'Heart Rate Variability Analysis using Correlation Dimension and Detrended Fluctuation Analysis', *Innovation and Research in BioMedical Engineering (ITBM-RBM)* **23**(6), 333–339.

Adhikari, R. & Agrawal, R. (2013), 'An Introductory Study on Time Series Modeling and Forecasting', *arXiv preprint arXiv:1302.6613* .

Aghaei, M., Dimiccoli, M., Ferrer, C. C. & Radeva, P. (2018), 'Towards Social Pattern Characterization in Egocentric Photo-streams', *Computer Vision and Image Understanding* **171**, 104–117.

Aghazadeh, O., Sullivan, J. & Carlsson, S. (2011), 'Novelty Detection from An Ego-centric Perspective', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 3297–3304.

Aguilar, E., Remeseiro, B., Bolax00F1os, M. & Radeva, P. (2018), 'Grab, Pay, and Eat: Semantic Food Detection for Smart Restaurants', *IEEE Transactions on Multimedia* **20**, 3266–3275.

Alatise, M. B. & Hancke, G. P. (2017), 'Pose Estimation of A Mobile Robot Based on Fusion of IMU Data and Vision Data Using An Extended Kalman Filter', *Sensors* **17**(10), 1–22.

Albatal, R., Gurrin, C., Zhou, J., Yang, Y., Carthy, D. & Li, N. (2013), 'Senseseer Mobile-cloud-based Lifelogging Framework', *IEEE International Symposium on Technology and Society (ISTAS)* pp. 144–146.

Androulakis, I. P. (2005), 'Selecting Maximally Informative Genes', *Computers & Chemical Engineering* **29**(3), 535–546.

Arcega, L., Font, J. & Cetina, C. (2013), 'Towards Memory-aware Services and Browsing through Lifelogging Sensing', *Sensors* **13**(11), 15113–15137.

Ashbrook, D., Lyons, K. & Clawson, J. (2006), 'Capturing Experiences Anytime, Anywhere', *IEEE Pervasive Computing* **5**(2), 8–11.

Askoxylakis, I., Brown, I., Dickman, P., Friedewald, M., Irion, K., Kosta, E., Langheinrich, M., McCarthy, P., Osimo, D., Papiotis, S. et al. (2011), 'To Log or not to Log? - Risks and Benefits of Emerging Life-logging applications', *Report on Enisa: European Union Agency for Network and Information Security* .

Atzori, L., Iera, A. & Morabito, G. (2010), 'The Internet of Things: A Survey', *Computer Networks* **54**(15), 2787–2805.

Ausloos, M. & Ivanova, K. (2000), 'Introducing False EUR and False EUR Exchange Rates', *Physica A: Statistical Mechanics and its Applications* **286**(1-2), 353–366.

Ausloos, M. & Ivanova, K. (2001), 'Power-law Correlations in the Southern-oscillation-index Fluctuations Characterizing El Niño', *Physical Review E* **63**(4), 047201.

Bell, G. & Gemmell, J. (2007), 'A Digital Life', *Scientific American* **296**(3), 58–65.

Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B. & Wood, K. (2007), 'The Use of a Wearable Camera, SenseCam, as A Pictorial Diary to Improve Autobiographical Memory in a Patient with Limbic Encephalitis: A Preliminary Report', *Neuropsychological Rehabilitation* **17**(4-5), 582–601.

Betancourt, A., Morerio, P., Regazzoni, C. S. & Rauterberg, M. (2015), 'The Evolution of First Person Vision Methods: A Survey', *IEEE Transactions on Circuits and Systems for Video Technology* **25**(5), 744–760.

Blum, M., Pentland, A. & Troster, G. (2006), 'Insense: Interest-Based Life Logging', *IEEE MultiMedia* **13**(4), 40–48.

Bolaños, M., Dimiccoli, M. & Radeva, P. (2017), 'Toward Storytelling from Visual Lifelogging: An Overview', *IEEE Transactions on Human-Machine Systems* **47**(1), 77–90.

Bolaños, M., Garolera, M. & Radeva, P. (2013), 'Active Labeling Application Applied to Food-Related Object Recognition', *The 5th International Workshop on Multimedia for Cooking & Eating Activities* pp. 45–50.

Bolaños, M., Garolera, M. & Radeva, P. (2014), 'Video Segmentation of Life-

logging Videos', *International Conference on Articulated Motion and Deformable Objects* pp. 1–9.

Bolaños, M., Peris, Á., Casacuberta, F., Soler, S. & Radeva, P. (2018), 'Egocentric video description based on temporally-linked sequences', *Journal of Visual Communication and Image Representation* **50**, 205–216.

Bouchaud, J.-P. & Potters, M. (2003), 'Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management', *Cambridge University Press* .

Box, G. E., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M. (2015), 'Time Series Analysis: Forecasting and Control', *John Wiley & Sons* .

Buckland, M. K. (1992), 'Emanuel Goldberg, Electronic Document Retrieval, and Vannevar Bush's Memex', *Journal of the American Society for Information Science* **43**(4), 284.

Buldyrev, S., Goldberger, A., Havlin, S., Peng, C., Stanley, H., Stanley, M. & Simons, M. (1993), 'Fractal Landscapes and Molecular Evolution: Modeling the Myosin Heavy Chain Gene Family', *Biophysical Journal* **65**(6), 2673–2679.

Burrus, C. S., Gopinath, R. A., Guo, H., Odegard, J. E. & Selesnick, I. W. (1998), 'Introduction to Wavelets and Wavelet Transforms: A Primer', *Prentice Hall New Jersey* .

Bush, V. & Think, A. W. M. (1945), 'As We May Think', *The atlantic monthly* **176**(1), 101–108.

Can, Z., Aslan, Z., Oguz, O. & Siddiqi, A. (2005), 'Wavelet Transforms of Meteorological Parameters and Gravity Waves', *Annales Geophysicae* **23**(3), 659–663.

Capobianco, E. (2004), 'Multiscale Analysis of Stock Index Return Volatility', *Computational Economics* **23**(3), 219–237.

Caraiani, P. (2012), 'Evidence of Multifractality from Emerging European Stock Markets', *PloS one* **7**(7), e40693.

Carpena, P., Bernaola-Galván, P., Coronado, A., Hackenberg, M. & Oliver, J. (2007), 'Identifying Characteristic Scales in the Human Genome', *Physical Review E* **75**(3), 032903.

Cartas, A., Marin, J., Radeva, P. & Dimiccoli, M. (2017), 'Recognizing Daily Activities from Egocentric Photo-Streams', *arXiv preprint:1710.04112* .

Castro, D., Hickson, S., Bettadapura, V., Thomaz, E., Abowd, G., Christensen, H. & Essa, I. (2015), 'Predicting Daily Activities from Egocentric Images using Deep Learning', *ACM International Symposium on Wearable Computers* pp. 75–82.

Chandrasekhar, V., Tan, C., Min, W., Liyuan, L., Xiaoli, L. & Hwee, L. J. (2014), 'Incremental Graph Clustering for Efficient Retrieval from Streaming Egocentric Video Data', *IEEE Conference on Pattern Recognition (ICPR)* pp. 2631–2636.

Chiu, B., Keogh, E. & Lonardi, S. (2003), 'Probabilistic Discovery of Time Series Motifs', *The 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* pp. 493–498.

Cochrane, J. H. (2005), 'Time Series for Macroeconomics and Finance', *Manuscript, University of Chicago* .

Conlon, T., Crane, M. & Ruskin, H. J. (2008), 'Wavelet Multiscale Analysis for Hedge Funds: Scaling and Strategies', *Physica A: Statistical Mechanics and its Applications* **387**(21), 5197–5204.

Conlon, T., Ruskin, H. J. & Crane, M. (2007), 'Random Matrix Theory and Fund of Funds Portfolio Optimisation', *Physica A: Statistical Mechanics and its applications* **382**(2), 565–576.

Conlon, T., Ruskin, H. J. & Crane, M. (2009), 'Multiscaled Cross-correlation Dynamics in Financial Time-series', *Advances in Complex Systems* **12**(04), 439–454.

Conrey, J. B., Farmer, D. W., Keating, J. P., Rubinstein, M. O. & Snaith, N. C. (2005), 'Integral Moments of L-functions', *London Mathematical Society* **91**(1), 33–104.

Cornish, C., Percival, D. & Bretherton, C. (2003), 'The WMTSA Wavelet Toolkit for Data Analysis in the Geosciences', *Conference on American Geophysical Union* pp. NG11A–0173.

Cornish, C. R., Bretherton, C. S. & Percival, D. B. (2006), 'Maximal Overlap Wavelet Statistical Analysis with Application to Atmospheric Turbulence', *Boundary-Layer Meteorology* **119**(2), 339–374.

Couillet, R. & Debbah, M. (2011), 'Random Matrix Methods for Wireless Communications', *Cambridge University Press* .

Danna, K. & Griffin, R. W. (1999), 'Health and Well-being in the Workplace: A Review and Synthesis of the Literature', *Journal of Management* **25**(3), 357–384.

Daubechies, I. (1992), 'Ten Lectures on Wavelets', *CBMS-NSF Regional Conference Series in Applied Mathematics* **61**.

De Jager, D., Wood, A. L., Merrett, G. V., Al-Hashimi, B. M., O'Hara, K., Shadbolt, N. R. & Hall, W. (2011), 'A Low-power, Distributed, Pervasive Healthcare System for Supporting Memory', *The 1st ACM MobiHoc Workshop on Pervasive Wireless Healthcare* p. 5.

Ding, H., Trajcevski, G., Scheuermann, P., Wang, X. & Keogh, E. (2008), 'Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures', *Very Large Database Endowment (VLDB Endowment)* **1**(2), 1542–1552.

Dobbins, C. & Fairclough, S. (2018), 'Signal Processing of Multimodal Mobile Lifelogging Data towards Detecting Stress in Real-World Driving', *IEEE Transactions on Mobile Computing* pp. 1–1.

Dodge, M. & Kitchin, R. (2007), ''Outlines of a World Coming into Existence': Pervasive Computing and the Ethics of Forgetting', *Environment and Planning B: Planning and Design* **34**(3), 431–445.

Doherty, A. R. (2009), 'Providing Effective Memory Retrieval Cues through Automatic Structuring and Augmentation of a Lifelog of Images', *PhD thesis, Dublin City University* .

Doherty, A. R., Caprani, N., Conaire, C. Ó., Kalnikaite, V., Gurrin, C., Smeaton, A. F. & O'Connor, N. E. (2011), 'Passively Recognising Human Activities through Lifelogging', *Computers in Human Behavior* **27**(5), 1948–1958.

Doherty, A. R., Hodges, S. E., King, A. C., Smeaton, A. F., Berry, E., Moulin, C. J., Lindley, S., Kelly, P. & Foster, C. (2013), 'Wearable Cameras in Health: the State of the Art and Future Possibilities', *American Journal of Preventive Medicine* **44**(3), 320–323.

Doherty, A. R., Ó Conaire, C., Blighe, M., Smeaton, A. F. & O'Connor, N. E. (2008), 'Combining Image Descriptors to Effectively Retrieve Events from Visual Lifelogs', *The 1st ACM International Conference on Multimedia Information Retrieval* pp. 10–17.

Doherty, A. R. & Smeaton, A. F. (2008), 'Automatically Segmenting Lifelog Data into Events', *The 9th International Workshop on Image Analysis for Multimedia Interactive Services* pp. 20–23.

Doherty, A. R., Smeaton, A. F., Lee, K. & Ellis, D. P. (2007), 'Multimodal Segmentation of Lifelog Data', *Conference on Large Scale Semantic Access to Content (Text, Image, Video, and Sound)* pp. 21–38.

Duan, W.-Q. & Stanley, H. E. (2011), 'Cross-correlation and the Predictability of Financial Return Series', *Physica A: Statistical Mechanics and its Applications* **390**(2), 290–296.

Dyson, F. J. (1962), 'Statistical Theory of the Energy Levels of Complex Systems. I', *Journal of Mathematical Physics* **3**(1), 140–156.

Dyson, F. J. & Mehta, M. L. (1963), 'Statistical Theory of the Energy Levels of Complex Systems. IV', *Journal of Mathematical Physics* **4**(5), 701–712.

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R. & Jermiin, L. S. (2017), 'Sensitivity and Specificity of Information Criteria', *PeerJ Preprints* **5**, e1103v3.

Eirola, E. & Lendasse, A. (2013), 'Gaussian Mixture Models for Time Series Modelling, Forecasting, and Interpolation', *International Symposium on Intelligent Data Analysis* pp. 162–173.

Fleck, R. & Fitzpatrick, G. (2009), 'Teachers' and Tutors' Social Reflection Around SenseCam Images', *International Journal of Human-Computer Studies* **67**(12), 1024–1036.

Fonseca, J. R. & Cardoso, M. G. (2007), 'Mixture-model Cluster Analysis Using Information Theoretical Criteria', *Intelligent Data Analysis* **11**(2), 155–173.

Foster, I., Zhao, Y., Raicu, I. & Lu, S. (2008), 'Cloud Computing and Grid Computing 360-degree Compared', *Grid Computing Environments Workshop* pp. 1–10.

Franklin, C. (2013), 'Gartner Predictions: 10 Ways IT Rocks the World (Again)', `http://www.enterpriseefficiency.com/mobile/author.asp?section_%20id=1129&doc_id=269098`. Accessed: 2014-12-10.

Galasso, F., Shankar Nagaraja, N., Jimenez Cardenas, T., Brox, T. & Schiele, B. (2013), A unified video segmentation benchmark: Annotation, metrics and analysis, *in* 'Proceedings of the IEEE International Conference on Computer Vision', pp. 3527–3534.

Gallegati, M. (2008), 'Wavelet Analysis of Stock Returns and Aggregate Economic Activity', *Computational Statistics & Data Analysis* **52**(6), 3061–3074.

Gemmell, J., Bell, G. & Lueder, R. (2006), 'MyLifeBits: A Personal Database for Everything', *Communications of the ACM* **49**(1), 88–95.

Gemmell, J., Bell, G., Lueder, R., Drucker, S. & Wong, C. (2002), 'MyLifeBits: Fulfilling the Memex Vision', *The 10th ACM International Conference on Multimedia* pp. 235–238.

Gençay, R., Selçuk, F. & Whitcher, B. (2001), 'Scaling Properties of Foreign Exchange Volatility', *Physica A: Statistical Mechanics and its Applications* **289**(1-2), 249–266.

Gifani, P., Rabiee, H., Hashemi, M., Taslimi, P. & Ghanbari, M. (2007), 'Optimal Fractal-scaling Analysis of Human EEG Dynamic for Depth of Anesthesia Quantification', *Journal of the Franklin Institute* **344**(3), 212–229.

Gopikrishnan, P., Plerou, V., Amaral, L. A. N., Meyer, M. & Stanley, H. E. (1999), 'Scaling of the Distribution of Fluctuations of Financial Market Indices', *Physical Review E* **60**(5), 5305.

Gopikrishnan, P., Rosenow, B., Plerou, V. & Stanley, H. E. (2000), 'Identifying Business Sectors from Stock Price Fluctuations', *arXiv preprint cond-mat/0011145* .

Gubbi, J., Buyya, R., Marusic, S. & Palaniswami, M. (2013), 'Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions', *Future Generation Computer Systems* **29**(7), 1645–1660.

Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L. & Albatal, R. (2016*a*), 'NTCIR Lifelog: The First Test Collection for Lifelog Research', *The 39th International ACM SIGIR conference on Research and Development in Information Retrieval* pp. 705–708.

Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L. & Albatal, R. (2016*b*), 'Overview of NTCIR-12 Lifelog Task', *NTCIR Conference on Evaluation of Information Access Technologies* pp. 354–360.

Gurrin, C., Smeaton, A. F., Doherty, A. R. et al. (2014), 'Lifelogging: Personal big data', *Foundations and Trends® in Information Retrieval* **8**(1), 1–125.

Hamilton, J. D. (1989), 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle', *Econometrica* **57**(2), 357–384.

Hammid, R., Maddi, S., Johnson, A., Bobick, A., Essa, I. & Isbell, C. L. (2005), 'Unsupervised Activity Discovery and Characterization from Event-streams', *The 21 Conference on Uncertainty in Artificial Intelligence* pp. 251–258.

Harper, R., Randall, D., Smyth, N., Evans, C., Heledd, L. & Moore, R. (2008), 'The Past is a Different Place: They Do Things Differently There', *The 7th ACM conference on Designing Interactive Systems* pp. 271–280.

Harper, R., Randall, D., Smythe, N., Evans, C., Heledd, L. & Moore, R. (2007), 'Thanks for the Memory', *The 21st British HCI Group Annual Conference on People and Computers* pp. 39–42.

Harvey, M., Langheinrich, M. & Ward, G. (2016), 'Remembering through

Lifelogging: A Survey of Human Memory Augmentation', *Pervasive and Mobile Computing* **27**, 14–26.

Harwood, J. (2009), 'Buckminster Fuller: Starting with the Universe', *Taylor & Francis* .

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N. & Wood, K. (2006), 'SenseCam: A Retrospective Memory Aid', *International Conference on Ubiquitous Computing* pp. 177–193.

Horenko, I., Dolaptchiev, S. I., Eliseev, A. V., Mokhov, I. I. & Klein, R. (2008), 'Metastable Decomposition of High-dimensional Meteorological Data with Gaps', *Journal of the Atmospheric Sciences* **65**(11), 3479–3496.

Hori, T. & Aizawa, K. (2003), 'Context-Based Video Retrieval System for the Life-log Applications', *The 5th ACM SIGMM International Workshop on Multimedia Information Retrieval* pp. 31–38.

Hu, F. (2016), 'Periodicity Detection and its Application in Lifelog Data', *PhD thesis, Dublin City University* .

Hu, F., Smeaton, A. F. & Newman, E. (2014), 'Periodicity Detection in Lifelog Data with Missing and Irregularly Sampled Data', *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* pp. 16–23.

Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R. & Stanley, H. E. (1999), 'Multifractality in Human Heartbeat Dynamics', *Nature* **399**(6735), 461.

Ivanov, P. C., Rosenblum, M. G., Peng, C.-K., Mietus, J., Havlin, S., Stanley,

H. E. & Goldberger, A. L. (1996), 'Scaling Behaviour of Heartbeat Intervals Obtained by Wavelet-based Time-series Analysis', *Nature* **383**(6598), 323.

Jánosi, I. M. & Müller, R. (2005), 'Empirical Mode Decomposition and Correlation Properties of Long Daily Ozone Records', *Physical Review E* **71**(5), 056126.

Jojic, N., Perina, A. & Murino, V. (2010), 'Structural Epitome: A Way to Summarize One's Visual Experience', *Conference on Advances in Neural Information Processing Systems* pp. 1027–1035.

Jothimani, D., Shankar, R. & Yadav, S. S. (2016), 'Discrete Wavelet Transform-based Prediction of Stock Index: A Study on National Stock Exchange Fifty index', *arXiv preprint arXiv:1605.07278* .

Kantelhardt, J. W. (2009), 'Fractal and Multifractal Time Series', *In Book Encyclopedia of Complexity and Systems Science* pp. 3754–3779.

Keogh, E., Chakrabarti, K., Pazzani, M. & Mehrotra, S. (2001), 'Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases', *Knowledge and Information Systems* **3**(3), 263–286.

Keogh, E. & Lin, J. (2005), 'Clustering of Time-series Subsequences is Meaningless: Implications for Previous and Future Research', *Knowledge and Information Systems* **8**(2), 154–177.

Kishida, K. & Kato, M. P. (2016), 'Overview of NTCIR-12', *NTCIR Conference on Evaluation of Information Access Technologies* pp. 1–7.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H. E., Goldreich, Y.

& Schellnhuber, H.-J. (1998), 'Indication of A Universal Persistence Law Governing Atmospheric Variability', *Physical Review Letters* **81**(3), 729.

Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. (2001), 'Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes', *Journal of molecular biology* **305**(3), 567–580.

Ladyman, J., Lambert, J. & Wiesner, K. (2013), 'What is a Complex System?', *European Journal for Philosophy of Science* **3**(1), 33–67.

Lagun, D., Ageev, M., Guo, Q. & Agichtein, E. (2014), 'Discovering Common Motifs in Cursor Movement Data for Improving Web Search', *The 7th ACM International Conference on Web Search and Data Mining* pp. 183–192.

Laloux, L., Cizeau, P., Bouchaud, J.-P. & Potters, M. (1999), 'Noise Dressing of Financial Correlation Matrices', *Physical Review Letters* **83**(7), 1467.

Laloux, L., Cizeau, P., Potters, M. & Bouchaud, J.-P. (2000), 'Random Matrix Theory and Financial Correlations', *International Journal of Theoretical and Applied Finance* **3**(03), 391–397.

Lämmer, S. & Helbing, D. (2008), 'Self-control of Traffic Lights and Vehicle Flows in Urban Road Networks', *Journal of Statistical Mechanics: Theory and Experiment* **2008**(04), 04019.

Lee, J.-M., Kim, D.-J., Kim, I.-Y., Park, K.-S. & Kim, S. I. (2002), 'Detrended Fluctuation Analysis of EEG in Sleep Apnea using MIT/BIH polysomnography data', *Computers in Biology and Medicine* **32**(1), 37–47.

Lee, M. L. & Dey, A. K. (2007), 'Providing Good Memory Cues for People with

Episodic Memory Impairment', *The 9th International ACM SIG ACCESS Conference on Computers and Accessibility* pp. 131–138.

Lee, P. A. & Ramakrishnan, T. (1985), 'Disordered Electronic Systems', *Reviews of Modern Physics* **57**(2), 287.

Li, N., Crane, M., Gurrin, C. & Ruskin, H. J. (2016), 'Finding Motifs in Large Personal Lifelogs', *Proceedings of the 7th Augmented Human International Conference* p. 9.

Li, N., Crane, M. & Ruskin, H. J. (2013), 'Automatically Detecting "Significant Events" on SenseCam', *International Journal of Wavelets, Multiresolution and Information Processing* **11**(06), 1350050.

Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2013*a*), 'Application of Statistical Physics for the Identification of Important Events in Visual Lifelogs', *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* pp. 589–592.

Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2013*b*), 'Multiscaled Cross-correlation Dynamics on SenseCam Lifelogged Images', *International Conference on Multimedia Modeling* pp. 490–501.

Li, N., Crane, M., Ruskin, H. J. & Gurrin, C. (2014), 'Random Matrix Ensembles of Time Correlation Matrices to Analyze Visual Lifelogs', *International Conference on Multimedia Modeling* pp. 400–411.

Li, N., Gurrin, C., Crane, M. & Ruskin, H. J. (2016), 'NTCIR-12 Lifelog Data Analytics', *Proceedings of the first Workshop on Lifelogging Tools and Applications* pp. 27–36.

Li, N. & Hopfgartner, F. (2016), 'To Log or Not To Log? SWOT Analysis of Self-tracking', *In Book Lifelogging* pp. 305–325.

Li, T., Choi, M., Fu, K. & Lin, L. (2018), 'Music Sequence Prediction with Mixture Hidden Markov Models', *arXiv preprint arXiv:1809.00842* .

Lin, J., Keogh, E., Lonardi, S. & Chiu, B. (2003), 'A Symbolic Representation of Time Series, with Implications for Streaming Algorithms', *The 8th ACM SIGMOD workshop on Research Issues in Data Mining and Knowledge Discovery* pp. 2–11.

Lin, J., Keogh, E., Lonardi, S. & Patel, P. (2002), 'Finding Motifs in Time Series', *In the 2nd Workshop on Temporal Data Mining, at the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* pp. 53–68.

Lin, J., Keogh, E., Wei, L. & Lonardi, S. (2007), 'Experiencing SAX: A Novel Symbolic Representation of Time Series', *Data Mining and Knowledge Discovery* **15**(2), 107–144.

Lin, W.-H. & Hauptmann, A. (2006), 'Structuring Continuous Video Recordings of Everyday Life Using Time-Constrained Clustering', *Multimedia Content Analysis, Management, and Retrieval* **6073**, 60730D.

Lindley, S. E., Marshall, C. C., Banks, R., Sellen, A. & Regan, T. (2013), 'Rethinking the Web as a Personal Archive', *The 22nd International Conference on World Wide Web* pp. 749–760.

Lou, J.-G., Lin, Q., Ding, R., Fu, Q., Zhang, D. & Xie, T. (2013), 'Software Analytics for Incident Management of Online Services: An Experience Report',

The 28th IEEE/ACM International Conference on Automated Software Engineering pp. 475–485.

Lu, Z. & Grauman, K. (2013), 'Story-Driven Summarization for Egocentric Video', *IEEE Conference on Computer Vision and Pattern Recognition* pp. 2714–2721.

Mann, S. (1997), 'Wearable Computing: A First Step Toward Personal Imaging', *Computer* **30**(2), 25–32.

Mann, S. (2004), 'Continuous Lifelong Capture of Personal Experience with EyeTap', *The the 1st ACM workshop on Continuous Archival and Retrieval of Personal Experiences* pp. 1–21.

Mann, S., Fung, J., Aimone, C., Sehgal, A. & Chen, D. (2005), 'Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences', *CHI Conference on Human Factors in Computing Systems* .

Martin, R. (2001), 'Noise Power Spectral Density Estimation based on Optimal Smoothing and Minimum Statistics', *IEEE Transactions on Speech and Audio Processing* **9**(5), 504–512.

Maurer, U., Smailagic, A., Siewiorek, D. P. & Deisher, M. (2006), 'Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions', *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)* pp. 4–116.

Mehta, M. L. (2004), 'Random Matrices', *Elsevier* .

Mehta, M. L. & Dyson, F. J. (1963), 'Statistical Theory of the Energy Levels of Complex Systems. V', *Journal of Mathematical Physics* **4**(5), 713–719.

Meyers, R. A. (2011), 'Mathematics of Complexity and Dynamical Systems', *Springer Science & Business Media* .

Mezzadri, F., Snaith, N. C., Hitchin, N. et al. (2005), 'Recent Perspectives in Random Matrix Theory and Number Theory', *Cambridge University Press* **322**.

Michael, K., Bowman, D., Jones, M. L. & Pringle, R. (2018), 'Robots and Socio-Ethical Implications', *IEEE Technology and Society Magazine* **37**(1), 19–21.

Minnen, D., Starner, T., Essa, I. A. & Isbell Jr, C. L. (2007), 'Improving Activity Discovery with Automatic Neighborhood Estimation', *International Joint Conferences on Artificial Intelligence Organization (IJCAI)* **7**, 2814–2819.

Minnen, D., Starner, T., Essa, I. & Isbell, C. (2006), 'Discovering Characteristic Actions from On-body Sensor Data', *The 10th IEEE International Symposium on Wearable Computers* pp. 11–18.

Moret, M. A., Zebende, G., Nogueira Jr, E. & Pereira, M. (2003), 'Fluctuation Analysis of Stellar X-ray Binary Systems', *Physical Review E* **68**(4), 041104.

Nevill-Manning, C. G., Wu, T. D. & Brutlag, D. L. (1998), 'Highly Specific Protein Sequence Motifs for Genome Analysis', *National Academy of Sciences* **95**(11), 5865–5871.

Newman, M. E. (2011), 'Complex Systems: A Survey', *arXiv preprint arXiv:1112.1440* .

Ordóñez, F. J. & Roggen, D. (2016), 'Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition', *Sensors* **16**(1), 1–25.

Patel, P., Bansal, D., Yuan, L., Murthy, A., Greenberg, A., Maltz, D. A., Kern, R., Kumar, H., Zikos, M., Wu, H. et al. (2013), 'Ananta: Cloud Scale Load Balancing', *ACM SIGCOMM Computer Communication Review* **43**(4), 207–218.

Pelletier, J. D. (1997), 'Analysis and Modeling of the Natural Variability of Climate', *Journal of Climate* **10**(6), 1331–1342.

Peng, C.-K., Buldyrev, S., Goldberger, A., Havlin, S., Simons, M. & Stanley, H. (1993), 'Finite-size Effects on Long-range Correlations: Implications for Analyzing DNA Sequences', *Physical Review E* **47**(5), 3730.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E. & Goldberger, A. L. (1994), 'Mosaic Organization of DNA Nucleotides', *Physical Review E* **49**(2), 1685.

Peng, C.-K., Havlin, S., Stanley, H. E. & Goldberger, A. L. (1995), 'Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **5**(1), 82–87.

Percival, D. B. & Walden, A. T. (2006), 'Wavelet Methods for Time Series Analysis', *Cambridge University Press* .

Perry, J. W., Kent, A. & Berry, M. M. (1955), 'Machine Literature Searching X. Machine Language; Factors Underlying its Design and Development', *American Documentation* **6**(4), 242–254.

Phinyomark, A., Phothisonothai, M., Limsakul, C. & Phukpattaranont, P. (2009), 'Detrended Fluctuation Analysis of Electromyography Signal to Identify Hand Movement', *The 2nd Biomedical Engineering International Conference* pp. 324–329.

Pitt, M. A. & Myung, I. J. (2002), 'When A Good Fit Can Be Bad', *Trends in Cognitive Sciences* **6**(10), 421–425.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. & Stanley, H. E. (2002), 'Random Matrix Approach to Cross Correlations in Financial Data', *Physical Review E* **65**(6), 066126.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N. & Stanley, H. E. (1999), 'Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series', *Physical Review Letters* **83**(7), 1471.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. N. & Stanley, H. E. (2000), 'A Random Matrix Theory Approach to Financial Cross-correlations', *Physica A: Statistical Mechanics and its Applications* **287**(3-4), 374–382.

Poleg, Y., Arora, C. & Peleg, S. (2014), 'Temporal Segmentation of Egocentric Videos', *IEEE Conference on Computer Vision and Pattern Recognition* pp. 2537–2544.

Pollock, D. (2006), 'Introduction to the Special Issue on Statistical Signal Extraction and Filtering', *Computational Statistics & Data Analysis* **50**(9), 2137–2145.

Putzig, L., Becherer, D. & Horenko, I. (2010), 'Optimal Allocation of a Futures Portfolio Utilizing Numerical Market Phase Detection', *SIAM Journal on Financial Mathematics* **1**(1), 752–779.

Qiu, Z., Gurrin, C., Doherty, A. R. & Smeaton, A. F. (2012), 'A Real-time Life Experience Logging Tool', *International Conference on Multimedia Modeling* pp. 636–638.

Reddy, S., Parker, A., Hyman, J., Burke, J., Estrin, D. & Hansen, M. (2007), 'Image Browsing, Processing, and Clustering for Participatory Sensing: Lessons from A DietSense Prototype', *The 4th Workshop on Embedded Networked Sensors* pp. 13–17.

Rodriguez, E., Echeverria, J. C. & Alvarez-Ramirez, J. (2007), 'Detrended Fluctuation Analysis of Heart Intrabeat Dynamics', *Physica A: Statistical Mechanics and its Applications* **384**(2), 429–438.

Rooksby, J., Rost, M., Morrison, A. & Chalmers, M. C. (2014), 'Personal Tracking as Lived Informatics', *The 32nd annual ACM conference on Human Factors in Computing Systems* pp. 1163–1172.

Ross, J. C., Vinutha, T. & Rao, P. (2012), 'Detecting Melodic Motifs from Audio for Hindustani Classical Music', *The 13th International Society for Music Information Retrieval Conference (ISMIR)* pp. 193–198.

Rubinov, M. & Sporns, O. (2010), 'Complex Network Measures of Brain Connectivity: Uses and Interpretations', *Neuroimage* **52**(3), 1059–1069.

Santhanam, M., Bandyopadhyay, J. N. & Angom, D. (2006), 'Quantum Spectrum as A Time Series: Fluctuation Measures', *Physical Review E* **73**(1), 015201.

Sarkar, S., Chatterjee, S. & Misra, S. (2018), 'Assessment of the Suitability of Fog Computing in the Context of Internet of Things', *IEEE Transactions on Cloud Computing* **6**(1), 46–59.

Schindler, K., Leung, H., Elger, C. E. & Lehnertz, K. (2006), 'Assessing Seizure Dynamics by Analysing the Correlation Structure of Multichannel Intracranial EEG', *Brain* **130**(1), 65–77.

Schütte, C. & Huisinga, W. (2003), 'Biomolecular Conformations Can Be Identified as Metastable Sets of Molecular Dynamics', *Handbook of Numerical Analysis* **10**, 699–744.

Selke, S. (2016), 'Lifelogging: Digital Self-tracking and Lifelogging-Between Disruptive Technology and Cultural Transformation', *Springer* .

Sellen, A. J., Fogg, A., Aitken, M., Hodges, S., Rother, C. & Wood, K. (2007), 'Do Life-logging Technologies Support Memory for The Past?: An Experimental Study Using SenseCam', *SIGCHI Conference on Human Factors in Computing Systems* pp. 81–90.

Sengupta, A. M. & Mitra, P. P. (1999), 'Distributions of Singular Values for Some Random Matrices', *Physical Review E* **60**(3), 3389.

Seo, Y., Choi, Y. & Choi, J. (2017), 'River Stage Modeling by Combining Maximal Overlap Discrete Wavelet Transform, Support Vector Machines and Genetic Algorithm', *Water* **9**(7), 525.

Siwy, Z., Ausloos, M. & Ivanova, K. (2002), 'Correlation Studies of Open and Closed State Fluctuations in An Ion Channel: Analysis of Ion Current through A Large-conductance Locust Potassium Channel', *Physical Review E* **65**(3), 031907.

Sornette, D. (2017), 'Why Stock Markets Crash: Critical Events in Complex Financial Systems', *Princeton University Press* .

Spector, A., Thorgrimsen, L., Woods, B., Royan, L., Davies, S., Butterworth, M. & Orrell, M. (2003), 'Efficacy of An Evidence-Based Cognitive Stimulation Therapy Programme for People with Dementia: Randomised Controlled Trial', *The British Journal of Psychiatry* **183**(3), 248–254.

Spriggs, E. H., De La Torre, F. & Hebert, M. (2009), 'Temporal Segmentation and Activity Classification from First-person Sensing', *IEEE Computer Society Conference On Computer Vision and Pattern Recognition Workshops* pp. 17–24.

Stam, C., Montez, T., Jones, B., Rombouts, S., Van Der Made, Y., Pijnenburg, Y. & Scheltens, P. (2005), 'Disturbed Fluctuations of Resting State EEG Synchronization in Alzheimer's Disease', *Clinical Neurophysiology* **116**(3), 708–715.

Stanley, H., Buldyrev, S., Goldberger, A., Havlin, S., Peng, C.-K. & Simons,

M. (1999), 'Scaling Features of Noncoding DNA', *Physica A: Statistical Mechanics and its Applications* **273**(1-2), 1–18.

Streams, K. (2012), 'EyeTap Digital Eye Glass Inventor Predicts the Future of Surveillance and Augmented Reality', `https://www.theverge.com/2012/11/5/3603162/steve-mann-eyetap-surveillance-sousveillance`. Accessed: 2018-12-03.

Sumpter, D. J. (2010), 'Collective Animal Behavior', *Princeton University Press* .

Talavera, E., Dimiccoli, M., Bolanos, M., Aghaei, M. & Radeva, P. (2015), 'R-clustering for Egocentric Video Segmentation', *Iberian Conference on Pattern Recognition and Image Analysis* pp. 327–336.

Talkner, P. & Weber, R. O. (2000), 'Power Spectrum and Detrended Fluctuation Analysis: Application to Daily Temperatures', *Physical Review E* **62**(1), 150.

Tan, C., Goh, H., Chandrasekhar, V., Li, L. & Lim, J.-H. (2014), Understanding the nature of first-person videos: Characterization and classification using low-level features, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops', pp. 535–542.

Tanaka, Y., Iwamoto, K. & Uehara, K. (2005), 'Discovery of Time-series Motif from Multi-dimensional Data based on MDL Principle', *Machine Learning* **58**(2-3), 269–300.

Tulino, A. M., Verdú, S. et al. (2004), 'Random Matrix Theory and Wire-

less Communications', *Foundations and Trends® in Communications and Information Theory* **1**(1), 1–182.

Ulfarsson, M. O. & Solo, V. (2008), 'Dimension Estimation in Noisy PCA with SURE and Random Matrix Theory', *IEEE Transactions on Signal Processing* **56**(12), 5804–5816.

Utsugi, A., Ino, K. & Oshikawa, M. (2004), 'Random Matrix Theory Analysis of Cross Correlations in Financial Markets', *Physical Review E* **70**(2), 026110.

Vandewalle, N. & Ausloos, M. (1998), 'Crossing of Two Mobile Averages: A Method for Measuring the Roughness Exponent', *Physical Review E* **58**(5), 6832.

Varela, F. J. & Coutinho, A. (1991), 'Second Generation Immune Networks', *Immunology Today* **12**(5), 159–166.

Vespier, U., Knobbe, A., Nijssen, S. & Vanschoren, J. (2012), 'MDL-Based Analysis of Time Series at Multiple Time-Scales', *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* pp. 371–386.

Wang, P. & Smeaton, A. F. (2012), 'Semantics-based Selection of Everyday Concepts in Visual Lifelogging', *International Journal of Multimedia Information Retrieval* **1**(2), 87–101.

Wang, P. & Smeaton, A. F. (2013), 'Using Visual Lifelogs to Automatically Characterize Everyday Activities', *Information Sciences* **230**, 147–161.

Wang, P., Sun, L., Smeaton, A. F., Gurrin, C. & Yang, S. (2018), 'Computer

Vision for Lifelogging: Characterizing Everyday Activities Based on Visual Semantics', *Computer Vision for Assistive Healthcare* pp. 249–282.

Wang, Z., Hoffman, M. D., Cook, P. R. & Li, K. (2006), 'VFerret: Content-based Similarity Search Tool for Continuous Archived Video', *ACM workshop on Continuous Archival and Retrieval of Personal Experiences* pp. 19–26.

Waters, D. & Garrett, J. (1996), 'Preserving Digital Information. Report of the Task Force on Archiving of Digital Information', *ERIC* .

Wiest, J., Höffken, M., Kreßel, U. & Dietmayer, K. (2012), 'Probabilistic Trajectory Prediction with Gaussian Mixture Models', *IEEE Intelligent Vehicles Symposium (IV)* pp. 141–146.

Wigner, E. P. (1951), 'On the Statistical Distribution of the Widths and Spacings of Nuclear Resonance Levels', *Cambridge Philosophical Society* **47**(4), 790–798.

Williams, B., Durvasula, P. & Brown, D. (1998), 'Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models', *Transportation Research Record: Journal of the Transportation Research Board* (1644), 132–141.

Yang, H., Pan, Z., Tao, Q. & Qiu, J. (2018), 'Online Learning for Vector Autoregressive Moving-average Time Series Prediction', *Neurocomputing* **315**, 9–17.

Yang, Y. (2005), 'Can The Strengths of AIC and BIC Be Shared? A Con-

flict Between Model Identification and Regression Estimation', *Biometrika* **92**(4), 937–950.

Ye, T. (2018), 'Visual Object Detection from Lifelogs using Visual Non-lifelog Data', *PhD thesis, Dublin City University* .

Yeh, C.-C. M., Kavantzas, N. & Keogh, E. (2017), 'Matrix Profile VI: Meaningful Multidimensional Motif Discovery', *IEEE International Conference on Data Mining (ICDM)* pp. 565–574.

Yeh, C.-C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., Silva, D. F., Mueen, A. & Keogh, E. (2016), 'Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View that includes Motifs, Discords and Shapelets', *IEEE 16th International Conference on Data Mining (ICDM)* pp. 1317–1322.

Zhang, H., Li, L., Jia, W., Fernstrom, J. D., Sclabassi, R. J. & Sun, M. (2010), 'Recognizing Physical Activity from Ego-Motion of a Camera', *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp. 5569–5572.

Zheng, X., Lian, Y. & Wang, Q. (2018), 'The Long-range Correlation and Evolution Law of Centennial-scale Temperatures in Northeast China', *PloS one* **13**(6), e0198238.

Zhu, X. (2008), 'The Application of Wavelet Transform in Digital Image Processing', *International Conference on MultiMedia and Information Technology* pp. 326–329.

Zubal', A. (2015), 'Application of Band Spectrum Regression in Economic Problems', *Master thesis, Univerzita Karlova* .